Motion of localized sources in general relativity: Gravitational self-force from quasilocal conservation laws

Marius Oltean, Richard J. Epp, Carlos F. Sopuerta, Alessandro D. A. M. Spallicci, Robert B. Mann

Motion of localized sources in general relativity: Gravitational self-force from quasilocal conservation laws

Marius Oltean,[1,2,3,4,5,*] Richard J. Epp,[6] Carlos F. Sopuerta,[1,2] Alessandro D. A. M. Spallicci,[4,5,7] and Robert B. Mann[6,8]

[1]*Institute of Space Sciences (ICE, CSIC), Campus Universitat Autònoma de Barcelona, Carrer de Can Magrans s/n, 08193 Cerdanyola del Vallès (Barcelona), Spain*
[2]*Institute of Space Studies of Catalonia (IEEC), Carrer del Gran Capità, 2-4, Edifici Nexus, despatx 201, 08034 Barcelona, Spain*
[3]*Departament de Física, Facultat de Ciències, Universitat Autònoma de Barcelona, Edifici C, 08193 Cerdanyola del Vallès (Barcelona), Spain*
[4]*Observatoire des Sciences de l'Univers en région Centre (OSUC), Université d'Orléans, 1A rue de la Férollerie, 45071 Orléans, France*
[5]*Laboratoire de Physique et Chimie de l'Environnement et de l'Espace (LPC2E), Centre National de la Recherche Scientifique (CNRS), 3A Avenue de la Recherche Scientifique, 45071 Orléans, France*
[6]*Department of Physics and Astronomy, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada*
[7]*Pôle de Physique, Collegium Sciences et Techniques (CoST), Université d'Orléans, Rue de Chartres, 45100 Orléans, France*
[8]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada*

(Received 21 November 2019; accepted 13 February 2020; published 26 March 2020)

An idealized "test" object in general relativity moves along a geodesic. However, if the object has a finite mass, this will create additional curvature in the spacetime, causing it to deviate from geodesic motion. If the mass is nonetheless sufficiently small, such an effect is usually treated perturbatively and is known as the gravitational self-force due to the object. This issue is still an open problem in gravitational physics today, motivated not only by basic foundational interest, but also by the need for its direct application in gravitational wave astronomy. In particular, the observation of extreme-mass-ratio inspirals by the future space-based detector LISA will rely crucially on an accurate modeling of the self-force driving the orbital evolution and gravitational wave emission of such systems. In this paper, we present a novel derivation, based on conservation laws, of the basic equations of motion for this problem. They are formulated with the use of a quasilocal (rather than matter) stress-energy-momentum tensor—in particular, the Brown-York tensor—so as to capture gravitational effects in the momentum flux of the object, including the self-force. Our formulation and resulting equations of motion are independent of the choice of the perturbative gauge. We show that, in addition to the usual gravitational self-force term, they also lead to an additional "self-pressure" force not found in previous analyses, and also that our results correctly recover known formulas under appropriate conditions. Our approach thus offers a fresh geometrical picture from which to understand the self-force fundamentally, and potentially useful new avenues for computing it practically.

DOI: 10.1103/PhysRevD.101.064060

I. INTRODUCTION

A. Gravitational waves and extreme-mass-ratio inspirals

The recent advent of gravitational wave astronomy—propelled by the ground-based direct detections achieved by the LIGO/Virgo Collaboration (see Ref. [1] for the detections during the O1 and O2 observing runs), the success of the LISA Pathfinder mission as a proof of principle for future space-based interferometric detectors [2,3], and the subsequent approval of the LISA mission for launch in the 2030s [4,5]—has brought a multitude of both practical and foundational problems to the foreground of gravitational physics today. While a plethora of possibilities for gravitational wave sources are actively being investigated theoretically and anticipated to become accessible observationally, both on the Earth as well as in space, the most ubiquitous class of such sources has manifestly been—and foreseeably will remain—the coalescence of compact object binaries [6,7]. These are two-body systems

*oltean@ice.cat

consisting of a pair of compact objects, say of masses M_1 and M_2, orbiting and eventually spiraling into each other. Each of these is, usually, either a *black hole* (BH) or a neutron star. There are also more general possibilities being investigated, including that of having a brown dwarf as one of the objects [8].

The LIGO/Virgo detections during the first scientific runs [1], O1 and O2, have all involved binaries of *stellar-mass compact objects* (SCOs) located in our local neighborhood. These have comparable masses, of the order of a few tens of solar masses each ($M_1 \sim M_2 \sim 10^{0-2} M_\odot$). In addition second- and third-generation terrestrial detectors can also eventually see *intermediate-mass-ratio inspirals*, binaries consisting of an intermediate-mass BH, of $10^{2-4} M_\odot$, and a SCO. While there is as yet no direct evidence for the existence of the former sorts of objects, there are good reasons to anticipate their detection (through gravitational waves) most likely at the centers of globular clusters, and their study provides an essential link to the strongly perturbative regime of compact object binary dynamics.

It is even further in this direction that future space-based detectors such as LISA are anticipated to take us. In particular, LISA is expected to see extreme-mass-ratio inspirals (EMRIs) [9], compact binaries where $M_1 \gg M_2$. An elementary sketch is depicted in Fig. 1. The more massive object could be a (*super*) *massive* black hole (MBH) of mass $M_1 = M \sim 10^{4-7} M_\odot$ located at a galactic center, with the significantly less massive object—effectively orbiting and eventually spiraling into the MBH—being a SCO: either a stellar-mass black hole or a neutron star, with $M_2 = m \sim 10^{0-2} M_\odot$.

Average estimates indicate that LISA will be able to see on the order of hundreds of EMRI events per year [10], with an expectation of observing, for each, thousands of orbital cycles over a period on the order of one year before the final plunge [11]. The trajectories defining these cycles and the gravitational wave signals produced by them will generally look much more complex than the relatively generic signals from mergers of stellar-mass black holes of comparable masses as observed, for example, by LIGO/Virgo.

EMRIs will therefore offer an ideal experimental milieu for strong gravity: the complicated motion of the SCO around the MBH will effectively "map out" the geometry—that is, the gravitational field—around the MBH, thus presenting us with an unprecedented opportunity for studying gravity in the very strong regime [10,12]. In particular, among the possibilities offered by EMRIs are the measurement of the mass and spin of the MBH to very high accuracy, testing the validity of the Kerr metric as the correct description of BHs within general relativity (GR), and testing GR itself as the correct theory of gravity.

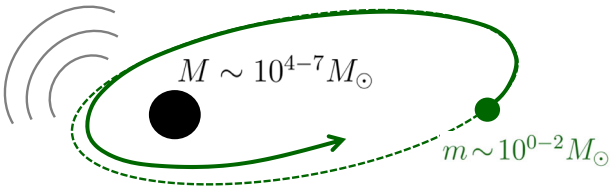

FIG. 1. Sketch of an EMRI, a two-body system consisting of a stellar-mass compact object (SCO), usually a stellar-mass black hole, of mass $m \sim 10^{0-2} M_\odot$, orbiting and eventually spiralling into a (super-) massive black hole (MBH), of mass $M \sim 10^{4-7} M_\odot$, and emitting gravitational waves in the process.

Yet, the richness of the observational opportunities presented by EMRIs comes with an inexorable cost: that is, a significant and as yet ongoing technical challenge in their theoretical modeling. This is all the more pressing as the EMRI signals expected from LISA are anticipated to be much weaker than the instrumental noise of the detector. Effectively, what this means is that extremely accurate models are necessary in order to produce the waveform templates that can be used to extract the relevant signals from the detector data stream. At the theoretical level, the problem of EMRI modeling cannot be tackled directly with numerical relativity (used for the LIGO/Virgo detections), simply due to the great discrepancy in (mass/length) scales; however, for the same reason, the approach that readily suggests itself is perturbation theory. See Fig. 2 for a graphic depicting the main methods used for compact object binary modeling in the different regimes. In particular, modeling the strong gravity, extreme-mass-ratio regime turns out to be equivalent to a general and quite old

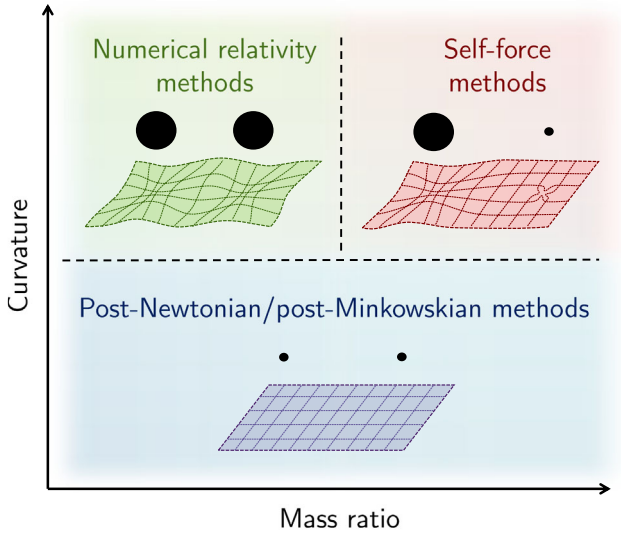

FIG. 2. The main approaches used in practice for the modeling of compact object binaries as a function of the mass ratio (increasing from 1) and the spacetime curvature involved. For low curvature (high separation between the bodies), post-Newtonian and post-Minkowskian methods are used. For high curvature (low separation) and low mass ratio, numerical relativity is used. For high curvature and extreme mass ratios, as the scale of a numerical grid would have to span orders of magnitude thus rendering it impracticable, perturbation theory must be used—in particular, self-force methods.

problem which can be posed in any (not just gravitational) classical field theory: the so-called *self-force* problem.

B. Self-force problem

Suppose we are dealing with a theory for a field $\psi(x)$ in some spacetime. If the theory admits a Lagrangian formulation, we can usually assume that the field equations have the general form

$$L[\psi(x)] = S(x), \tag{1}$$

where L is a (partial, possibly nonlinear and typically second-order) differential operator, and we refer to S as the *source* of the field ψ. Broadly speaking, the problem of the self-force is to find solutions $\psi(x)$ satisfying (1) when S is "localized" in spacetime. Intuitively, it is the question of how the existence of a dynamical (field-generating) "small object" (a mass, a charge, etc.) backreacts upon the total field ψ, and hence in turn upon its own future evolution subject to that field. Thus, an essential part of any detailed self-force analysis is a precise specification of what exactly it means for S to be localized. In standard approaches, one typically devises a perturbative procedure whereby S ends up being approximated as a distribution, usually a Dirac delta, compactly supported on a worldline—that is, the *background* (zeroth perturbative order) worldline of the small object. However, this already introduces a nontrivial mathematical issue: if L is nonlinear (in the standard partial differential equation sense), then the problem (1) with a distributional source S is mathematically ill defined, at least within the classical theory of distributions [13] where products of distributions do not make sense [14].[1]

One might therefore worry that nonlinear physical theories, such as GR, would *a priori* not admit solutions sourced by distributions, and we refer the interested reader to Ref. [20] for a classic detailed discussion of this topic. The saving point is that, while the full field equation (in this case, the Einstein equation) may indeed be generally nonlinear, if we devise a perturbative procedure (where the meaning of the perturbation is prescribed in such a way as to account for the presence of the small object itself), then the first-order field equation is, by construction, linear in the (first-order) perturbation $\delta\psi$ of ψ. Thus, assuming the background field is a known exact solution of the theory, it always makes sense to seek solutions $\delta\psi$ to

$$\delta L[\delta\psi(x)] = S(x), \tag{2}$$

for a distributional source S, where δL indicates the first-order part of the operator L in the full field equation (1). As this only makes sense for the (linear) first-order problem, such an approach becomes again ill defined if we begin to ask about the (nonlinear) second- or any higher-order problem. Additional technical constructions are needed to deal with these, the most common of which for the gravitational self-force has been the so-called "puncture" (or "effective source") method [11,21–23]; similar ideas have proven to be very useful also in numerical relativity [24,25]. For work on the second-order equation of motion for the gravitational self-force problem, see e.g., Refs. [26–28]. For now, we assume that we are interested here in the first-order self-force problem (2) only.

Now, concretely, in GR, our physical field ψ is simply the spacetime metric g_{ab} (where latin letters from the beginning of the alphabet indicate spacetime indices), and following standard convention, we denote a first-order perturbation thereof by $\delta g_{ab} = h_{ab}$. The problem (2) is then just the first-order Einstein equation,

$$\delta G_{ab}[h_{cd}] = \kappa T^{\mathrm{PP}}_{ab}, \tag{3}$$

where G_{ab} is the Einstein tensor, $\kappa = 8\pi$ (in geometrized units $c = G = 1$) is the Einstein constant, and T^{PP}_{ab} is the energy-momentum tensor of a "point particle" (PP) compactly supported on a given worldline $\mathring{\mathscr{C}}$. We will return later to discussing this more precisely, but in typical approaches, $\mathring{\mathscr{C}}$ turns out to be a geodesic—that is, the "background motion" of the small object, which is in this case a small mass.[2] Thus, simply solving (3) for h_{ab} assuming a fixed $\mathring{\mathscr{C}}$ for all time, though mathematically well defined, is by itself physically meaningless; it would simply give us the metric perturbations caused by a small object eternally moving on the same geodesic. Instead, what we would ultimately like is a way to take into account how h_{ab} modifies the motion of the small object itself. Thus, in addition to the field equation (3), any self-force analysis must be supplemented by an *equation of motion* (EoM) telling us, essentially, how to move from a given

[1]Nonlinear theories of distributions are being actively investigated by mathematicians [15–17]. Some work has been done to apply these to the electromagnetic self-force problem [18] and to study their general applicability in GR [19]; however, at this point, to our knowledge, their potential usefulness for the gravitational self-force problem has not been contemplated to any significant extent.

[2]The problem of deriving geodesic motion for appropriately defined nondynamical "test particles" from the Einstein equation (in lieu of postulating it as an independent axiom of GR additional to the Einstein equation) is a long-standing and interesting issue in its own right. Einstein was involved in some of the earliest work on this [29], and over the decades, various proofs have been put forward outside of the context of the gravitational self-force problem. See Refs. [30,31] for some of the most famous such proofs. See also Ref. [32] for a recent general review of the most widely used approaches as well as an interesting novel proposal. We consider later in this paper in detail one approach to the gravitational self-force which also proves geodesic motion as the background motion of point particles in GR.

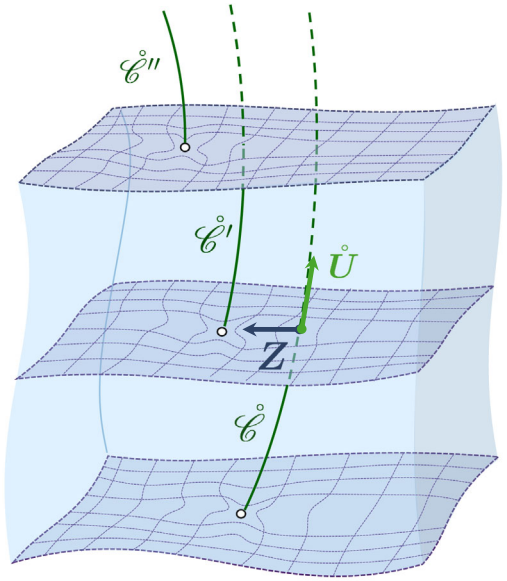

FIG. 3. A depiction of the perturbative problem for the GSF. In particular, this represents one of the most popular conceptions of a so-called self-consistent approach [33]: at a given step (on a given Cauchy surface) in the time evolution problem, one computes the "correction to the motion" away from geodesic ($\mathring{\mathscr{C}}$) in the form of a deviation vector Z^a, determined by the GSF. Then, at the next time step, one begins on a new ("corrected") geodesic ($\mathring{\mathscr{C}}'$), computes the new deviation vector, and so on.

background geodesic $\mathring{\mathscr{C}}$ at one step in the (ultimately numerical) time evolution problem to a new background geodesic $\mathring{\mathscr{C}}'$ at the next time step—with respect to which the field equation (3) is solved anew, and so on. This is sometimes referred to as a "self-consistent" approach. See Fig. 3 for a visual depiction.

The first proposal for an EoM for the *gravitational self-force* (GSF) problem was put forward in the late 1990s, since known as the MiSaTaQuWa equation after its authors [34,35]. On any $\mathring{\mathscr{C}}$, it reads

$$\ddot{Z}^a = -\mathring{E}_b{}^a Z^b + F^a[h^{\rm tail}_{cd}; \mathring{U}^e]. \tag{4}$$

The lhs is a second (proper) time derivative of a *deviation vector* Z^a on $\mathring{\mathscr{C}}$ pointing in the direction of the "true motion" (away from $\mathring{\mathscr{C}}$), to be defined more precisely later. On the rhs, $\mathring{E}_{ab}$ is the electric part of the Weyl tensor on $\mathring{\mathscr{C}}$, such that the first term is a usual "geodesic deviation" term. The second term on the rhs is the one usually understood as being responsible for self-force effects: $F^a[\cdot;\cdot]$ is a four-vector functional of a symmetric rank-2 contravariant tensor and a vector, to which we refer in general (for any arguments) as the *GSF functional*. In any spacetime with a given metric $\mathring{g}_{ab}$ and compatible derivative operator $\mathring{\nabla}_a$, it is explicitly given by the following simple action of a first-order differential operator:

$$F^a[H_{bc}; V^d] = -(\mathring{g}^{ab} + V^a V^b)\left(\mathring{\nabla}_c H_{bd} - \frac{1}{2}\mathring{\nabla}_b H_{cd}\right) V^c V^d. \tag{5}$$

While this is easy enough to calculate once one knows the arguments, the main technical challenge in using the MiSaTaQuWa equation (4) lies precisely in the determination thereof: in particular, $h^{\rm tail}_{ab}$ is not the *full* metric perturbation h_{ab} which solves the field equation (3), but instead represents what is called the "tail" integral of the Green functions of h_{ab} [36]. This quantity is well defined, but difficult to calculate in practice and usually requires the fixing of a perturbative gauge—typically the *Lorenz* gauge, $\mathring{\nabla}^b(h_{ab} - \frac{1}{2}h_{cd}\mathring{g}^{cd}\mathring{g}_{ab}) = 0$. Physically, $h^{\rm tail}_{ab}$ can be thought of as the part of the full perturbation h_{ab} which is scattered back by the spacetime curvature. (In this way, h_{ab} can be regarded as the sum of $h^{\rm tail}_{ab}$ and the remainder, what is sometimes called the "instantaneous" or "direct" part $h^{\rm direct}_{ab}$, responsible for waves radiated to infinity [37].)

An alternative, equivalent GSF EoM was proposed by Detweiler and Whiting in the early 2000s [38]. It relies upon a regularization procedure for the metric perturbations, i.e., a choice of a decomposition for h_{ab} [the full solution of the field equation (3)] into the sum of two parts: one which diverges—in fact, one which contains *all* divergent contributions—on $\mathring{\mathscr{C}}$, denoted $h^{\rm S}_{ab}$ (the so-called singular field, related to the direct part of the metric perturbation), and a remainder which is finite, $h^{\rm R}_{ab}$ (the so-called regular field, related to the tail part), so that one writes $h_{ab} = h^{\rm S}_{ab} + h^{\rm R}_{ab}$. An analogy with the self-force problem in electromagnetism gives some physical intuition behind how to interpret the meaning of this decomposition [11], with $h^{\rm S}_{ab} \sim m/r$ having the heuristic form of a "Coulombian self-field." However, no procedure is known for obtaining the precise expression of $h^{\rm S}_{ab}$ in an arbitrary perturbative gauge, and moreover, once a gauge is fixed (again, usually the Lorenz gauge), this splitting is not unique [11]. Nevertheless, if and when such an $h^{\rm S}_{ab}$ is obtained (from which we thus also get $h^{\rm R}_{ab} = h_{ab} - h^{\rm S}_{ab}$), the Detweiler-Whiting EoM for the GSF reads

$$\ddot{Z}^a = -\mathring{E}_b{}^a Z^b + F^a[h^{\rm R}_{cd}; \mathring{U}^e]. \tag{6}$$

The EoM's (4) and (6) are equivalent in the Lorenz gauge and form the basis of the two most popular methods used today for the numerical computation of the GSF. Yet a great deal of additional technical machinery is required for handling gauge transformations. This is essential because,

in the EMRI problem, the background spacetime metric—that of the MBH—is usually assumed to be Schwarzschild-Droste or Kerr. Perturbation theory for such spacetimes has been developed and is most easily carried out in, respectively, the so-called Regge-Wheeler and radiation gauges; in other words, in practice, it is often difficult (though not infeasible—see, e.g., Ref. [39]) to compute h_{ab} directly in the Lorenz gauge for use in (4) or (6).

A proposal for an EoM for the GSF problem that is valid in a wider class of perturbative gauges was presented by Gralla in 2011 [40]. It was therein formulated in what are called "parity-regular" gauges, i.e., gauges satisfying a certain parity condition. This condition ultimately has its origins in the Hamiltonian analysis of Regge and Teitleboim in the 1970s [41], wherein the authors introduce it in order to facilitate the vanishing of certain surface integrals and thus to render certain general-relativistic Hamiltonian notions, such as multipoles and "center of mass," well defined mathematically. In parity-regular gauges (satisfying the Regge-Teitleboim parity condition), the Gralla EoM—mathematically equivalent, in the Lorenz gauge, to the MiSaTaQuWa and the Detweiler-Whiting EoM's—is

$$\ddot{Z}^a = -\mathring{E}_b{}^a Z^b + \frac{1}{4\pi}\lim_{r\to 0}\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} F^a[h_{cd}; \mathring{U}^e]. \tag{7}$$

The GSF (last) term on the rhs is obtained in this approach by essentially relating the deviation vector Z^a (the evolution of which is expressed by the lhs) with a gauge transformation vector and then performing an "angle average" over a "small" r-radius two-sphere $\mathbb{S}^2_r$, with $\boldsymbol{\epsilon}_{\mathbb{S}^2}$ the volume form of the unit two-sphere, of the so-called bare GSF, $F^a[h_{bc}; \mathring{U}^d]$. The latter is just the GSF functional [Eq. (5)] evaluated directly using the full metric perturbatiuon h_{ab} (i.e., the tail plus direct parts, or equivalently the "regular" plus "singular" parts), around (rather than at the location of) the distributional source. The point therefore is that this formula never requires the evaluation of h_{ab} on $\mathring{\mathscr{C}}$ itself, where it is divergent by construction; instead, away from $\mathring{\mathscr{C}}$ it is always finite,[3] and (7) says that it suffices to compute the GSF functional (5) with h_{ab} directly in the argument (requiring no further transformations), and integrate it over a small sphere.

[3]This is true provided one does not transform to a perturbative gauge wherein any of the multipole moments of $\boldsymbol{h}$ diverge away from $\mathring{\mathscr{C}}$ as a consequence of the gauge definition. Examples of gauges leading to such divergences are the "half-string" and "full-string" radiation gauges of Ref. [42], which exhibit stringlike singularities in $\boldsymbol{h}$ along a radial direction. Nevertheless, in this work, it was also shown that one can define a no-string radiation gauge which is in the parity-regular class, and where the singularities in $\boldsymbol{h}$ remain only on $\mathring{\mathscr{C}}$, thus rendering the integral (5) well defined.

The manifest advantage of (7) relative to (4) or (6) is that no computations of tail integrals or regularizations of the metric perturbations are needed at all. Yet, to our knowledge, there has thus far been no attempted numerical computation of the GSF using this formula. One of the issues with this remains that of the perturbative gauge: depending upon the detailed setup of the problem, one may still not easily be able to compute h_{ab} directly in a parity-regular gauge (although manifestly, working in the parity-regular "no-string" radiation gauge [42] may be useful for a GSF calculation in Kerr), i.e., a gauge in which (7) is strictly valid, and so further gauge transformations may be needed. Aside from the practical issues with a possible numerical implementation of this, there is also a conceptual issue: this formula originates from an essentially mathematical argument—by a convenient "averaging" over the angles—so as to make it well defined in a Hamiltonian setting via a relation to a canonical definition of the center of mass. Yet its general form as a closed two-surface integral suggestively hints at the possibility of interpreting it not merely as a convenient mathematical relation, but as a real physical flux of (some notion of) "gravitational momentum." We contend and will demonstrate in this paper that indeed an even more general version of (7), not restricted by any perturbative gauge choice (so long as one does not construct it in such a way that produces divergences in $\boldsymbol{h}$ away from $\mathring{\mathscr{C}}$), results from the consideration of momentum conservation laws in GR.

C. Self-force problem via conservation laws

The idea of using conservation laws for tackling the self-force problem was appreciated and promptly exploited quite early on for the electromagnetic self-force. In the 1930s [43], Dirac was the first to put forward such an analysis in flat spacetime, and later on in 1960 [44], DeWitt and Brehme extended it to nondynamically curved spacetimes.[4] In such approaches, it can be shown[5] that the EoM for the electromagnetic self-force follows from local conservation expressions of the form

$$\Delta P^a = \int_{\Delta\mathscr{B}} \boldsymbol{\epsilon}_{\mathscr{B}} T^{ab} n_b, \tag{8}$$

where the lhs expresses the flux of *matter* four-momentum P^a (associated with T_{ab}) between the "caps" of (i.e., closed spatial two-surfaces delimiting) a portion (or "time interval") of a thin worldtube boundary $\mathscr{B}$ (topologically $\mathbb{R}\times\mathbb{S}^2$), with natural volume form $\boldsymbol{\epsilon}_{\mathscr{B}}$ and (outward-directed) unit normal n^a (see Fig. 4). In particular, one takes a time derivative of (8) to obtain an EoM expressing

[4]By this, we mean spacetimes with nonflat but fixed metrics, which do not evolve dynamically (gravitationally) in response to the matter stress-energy-momentum present therein.

[5]See Ref. [45] for a basic and more contemporary presentation.

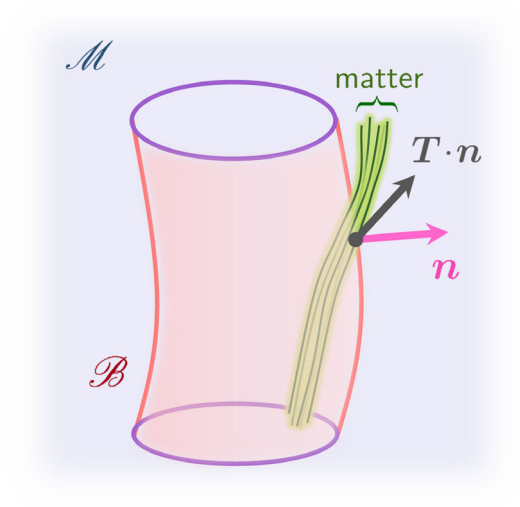

FIG. 4. A worldtube boundary $\mathscr{B}$ (topologically $\mathbb{R}\times\mathbb{S}^2$) in $\mathscr{M}$, with (outward-directed) unit normal n^a. The change in matter four-momentum between two constant time slices of this worldtube is given by the flux of the normal projection (in one index) of the matter stress-energy-momentum tensor T_{ab} through the portion of $\mathscr{B}$ bounded thereby.

the time rate of change of momentum in the form of a closed spatial two-surface integral (by differentiating the worldtube boundary integral). For the electromagnetic self-force problem, the introduction of an appropriate matter stress-energy-momentum tensor T_{ab} into Eq. (8) and a bit of subsequent argumentation reduces the integral expression to the famous Lorentz-Dirac equation; on a spatial three-slice in a Lorentz frame and in the absence of external forces, for example, this simply reduces to $\dot{P}^i=\frac{2}{3}q^2\dot{a}^i$ for a charge q.

Along these lines, see also Ref. [46] for a treatment of the scalar and electromagnetic self-forces based on the notion of a limit of an extended charge. Formulations of the scalar and electromagnetic self-forces using generalized Killing fields have more recently been put forward in Refs. [47,48]. Then in Ref. [49], the idea of generalized Killing fields was also used to study the gravitational self-force problem by employing definitions of momenta as bulk (volume) integrals (over the worldtube interior), involving the internal matter stress-energy-momentum tensor of the moving object (mass). This in fact proved the validity of the Detweiler-Whiting approach, i.e., viewing the "point mass" as moving on a geodesic in an effective (regularized) metric, to all multiple orders and also for the angular momentum. As we will shortly elaborate, we posit that gravitational momentum in GR cannot fundamentally be regarded as a bulk (local) density, and a more detailed analysis to reveal how the (locally formulated) results of Ref. [49] relate to our work here would be very interesting.

The success of conservation law approaches for formulating the electromagnetic self-force in itself inspires hope that the same may be done in the case of the GSF problem. In particular, Gralla's EoM (7) strongly hints at the possibility of understanding the rhs not just as a mathematical (angle averaging) device, but as a *true, physical flux of gravitational momentum* arising from a consideration of conservation expressions.

Nevertheless, to our knowledge, there has thus far been no proposed general treatment of the GSF following such an approach. This may, in large part, be conceivably attributed to the notorious conceptual difficulties surrounding the very question of the basic formulation of conservation laws in GR. Local conservation laws, along the lines of Eq. (8) that can readily be used for electromagnetism, no longer make sense fundamentally once gravity is treated as dynamical. The reason has a simple explanation in the equivalence principle [50]: one can always find a local frame of reference with a vanishing local "gravitational field" (metric connection coefficients), and hence a vanishing local "gravitational energy-momentum," irrespective of how one might feel inclined to define the latter.[6]

A wide variety of approaches have been taken over the decades toward formulating sensible notions of gravitational energy-momentum, with still no general consensus among relativists today on which to qualify as "the best" [52,53]. Often the preference for employing certain definitions over others may simply come down to context or convenience, but in any case, there exist agreements between the most typical definitions in various limits. A very common feature among them is the idea of replacing a local notion of gravitational energy-momentum, i.e., energy-momentum as a volume density, with what is known as a *quasilocal* energy-momentum, i.e., energy-momentum as a boundary density. The typical Hamiltonian definitions of the (total) gravitational energy-momentum for an asymptotically flat spacetime, for example, are of such a form. Among the most commonly used generalizations of these definitions to arbitrary (finite) spacetime

[6]It is worth remarking here that, in a perturbative setting, an approach that is sometimes taken is to work with an "effective" local gravitational stress-energy-momentum tensor, defined as the rhs of a suitably rearranged (first-order) Einstein equation. This is a common tactic often used for studying, for example, the energy-momentum of gravitational waves, with some applicable caveats (see, e.g., Chapter 35 of Ref. [50]). In fact, one of the first formulations of the gravitational self-force—in particular, the derivation of the MiSaTaQuWa EoM (4) presented in Sec. III of Ref. [34]—made use of the local conservation (vanishing of the spacetime divergence) of a suitably defined local tensor of such a sort (in analogy with the approach of DeWitt and Brehme to the electromagnetic self-force [44]). We elaborate in the remainder of this subsection and at greater length in Sec. II on why such a notion of gravitational conservation principles, while demonstrably useful for operational computations in some situations, cannot in general be expected to capture the fundamentally quasilocal (boundary density) nature of gravitational energy-momentum. See Ref. [51] for a detailed discussion and comparison between these two (local and quasilocal) views of gravitational energy-momentum.

regions was proposed in the early 1990s by Brown and York [54], and follow from what is now eponymously known as the Brown-York stress-energy-momentum tensor. It is a quasilocal tensor, meaning it is only defined on the boundary of an arbitrary spacetime region. For example, using this, the total (matter plus gravitational) energy inside a spatial volume is given up to a constant factor by the closed two-surface (boundary) integral of the trace of the boundary extrinsic curvature—precisely in agreement with the Hamiltonian definition of energy for the entire spacetime in the appropriate limit (where the closed two-surface approaches a two-sphere at asymptotically flat spatial infinity) but, in principle, applicable to any region in any spacetime.

The formulation of general energy-momentum conservation laws in GR from the Brown-York tensor has been achieved in recent years with the use of a construction called *quasilocal frames* [51], a concept first proposed in Ref. [55]. Essentially, the idea is that, in order to describe a (finite) gravitational system of interest and its associated boundary fluxes of gravitational energy-momentum, it does not suffice to merely specify, as in the local matter conservation laws of the form of Eq. (8), a worldtube boundary $\mathscr{B}$ (the interior of which is to be regarded as containing that system, and across which its energy-momentum flux is measured) as an embedded submanifold of $\mathscr{M}$. What is in fact required is the specification of a *congruence* making up this worldtube boundary, i.e., a two-parameter family of timelike worldlines with some chosen four-velocity field representing the motion of a topological two-sphere's worth of *quasilocal observers*. We will motivate this construction in greater amplitude shortly, but the reason for needing it is basically to be able to meaningfully define "time-time" and "time-space" directions on $\mathscr{B}$ for our conservation laws. A congruence of this sort is what is meant by a quasilocal frame.

The enormous advantage in using these quasilocal conservation laws over other approaches lies in the fact that they hold in any arbitrary spacetime. Thus, the existence of Killing vector fields—a typical requirement in other conservation law formulations—is in no way needed here.

This idea has been used successfully in a number of applications so far [51,56–61]. These include the resolution of a variation of Bell's spaceship paradox[7] in which a box accelerates rigidly in a transverse, uniform electric field [56], recovering under appropriate conditions the typical (but more limited) local matter conservation expressions of the form of Eq. (8) from the quasilocal ones [51] and application to post-Newtonian theory [58] and to relativistic geodesy [60,61].

A similar idea to quasilocal frames, called *gravitational screens*, was proposed more recently in Refs. [64,65]. There, the authors also make use of quasilocal ideas to develop conservation laws very similar in style and form to those obtained via quasilocal frames. A detailed comparison between these two approaches has thus far not been carried out, but it would be very interesting to do so in future work. In particular, the notion of gravitational screens has been motivated more from thermodynamic considerations, and similarly casting quasilocal frames in this language could prove quite fruitful. For example, just as these approaches have given us operational definitions of concepts like the "energy-momentum in an arbitrary spacetime region" (and not just for special cases such as an entire spacetime), they may help to do the same for concepts like "entropy in an arbitrary spacetime region" (and not just for known special cases such as a black hole).

D. Executive summary of the paper

We now summarize the structure and main result of this paper.

Section II is entirely devoted to an overview of quasilocal frames and quasilocal conservation laws, in complete self-contained technical detail for our purposes here.

In Sec. III, we prove the main general result of this paper from the quasilocal momentum conservation law: in particular, we show that for any localized gravitational system, the change between some initial and final time in its total linear momentum[8] p,

$$\Delta \mathrm{p}^{(\boldsymbol{\phi})} = \mathrm{p}^{(\boldsymbol{\phi})}_{\text{final}} - \mathrm{p}^{(\boldsymbol{\phi})}_{\text{initial}}, \tag{9}$$

along a spatial direction determined by a vector $\boldsymbol{\phi}$ (the precise meaning of which is to be elaborated later), is given by the following flux through the worldtube boundary interval $\Delta\mathscr{B}$ bounded by these times (see Fig. 4):

$$\Delta \mathrm{p}^{(\boldsymbol{\phi})} = -\frac{1}{4\pi}\int_{\Delta\mathscr{B}} \boldsymbol{\epsilon}_{\mathscr{B}}\, \frac{1}{r}\boldsymbol{\phi}\cdot\boldsymbol{\mathcal{F}}[h_{ab}; \mathring{u}^c] + \mathcal{O}(r). \tag{10}$$

Here, $\mathscr{F}^a[\cdot;\cdot]$ is an *extended* GSF functional. In particular, it is the usual GSF functional F^a [see Eq. (5)] plus a novel piece to which we refer as the *gravitational self-pressure force* $\wp^a$, arising from a quasilocal pressure effect (also to be elaborated later):

$$\mathscr{F}^a = F^a + \wp^a. \tag{11}$$

The first argument of this functional $\mathscr{F}^a$, just as in Gralla's formula [our Eq. (7)], is the metric perturbation h_{ab} on $\mathscr{B}$. This avoids any potential singularities in inside the worldtube, i.e., within the spacetime volume of which $\mathscr{B}$ is the (exterior) boundary, and therefore the need for performing regularizations or any further transformations. The second

[7]Proposed initially by E. Dewan and M. Beran [62] and later made popular by J.S. Bell's version [63].

[8]This is written using lowercase typewriter font (p) to distinguish it from the purely matter four-momentum P^a associated with T_{ab}.

argument of $\mathscr{F}^a$ in our result [see Eq. (10)], i.e., $\overset{\circ}{u}{}^a$, is the four-velocity *not* of any background geodesic contained inside the worldtube, as in the typical GSF EoM's discussed earlier—indeed, strictly speaking, the main result [Eq. (10)] holds without necessarily having to even introduce any such geodesic, or more generally, without having to say anything specific about the content of the worldtube interior—but instead that of the background quasilocal observers on $\mathscr{B}$ itself.

Manifestly, our formula [Eq. (10)] bears significant resemblance to that of Gralla [Eq. (7)], and it is the scope of Sec. IV to compare the two in the appropriate setting. For this, we introduce the general setup of the Gralla-Wald approach to the GSF, and apply our conservation law formula for two choices of quasilocal frames: one which is inertial with the SCO in the perturbed spacetime (and hence not inertial with the geodesic-following point particle in the background) and one which is inertial with the geodesic-following point particle in the background (and hence not inertial with the SCO in the perturbed spacetime). We derive the EoM's in both of these cases and discuss their correspondence to the known GSF EoM's.

Finally, Sec. V concludes the paper.

E. Notation and conventions

We work in the $(-,+,+,+)$ signature of spacetime. Script uppercase letters $(\mathscr{A},\mathscr{B},\mathscr{C},\ldots)$ are reserved for denoting mathematical spaces (manifolds, curves, etc.). The n-dimensional Euclidean space is denoted as usual by $\mathbb{R}^n$, the n-sphere of radius r is denoted by $\mathbb{S}^n_r$, and the unit n-sphere is denoted by $\mathbb{S}^n = \mathbb{S}^n_1$. For any two spaces $\mathscr{A}$ and $\mathscr{B}$ that are topologically equivalent (i.e., homeomorphic), we indicate this by writing $\mathscr{A} \simeq \mathscr{B}$.

The set of (k,l)-tensors on any manifold $\mathscr{U}$ is denoted by $\mathscr{T}^k{}_l(\mathscr{U})$. In particular, $T\mathscr{U} = \mathscr{T}^1{}_0(\mathscr{U})$ is the tangent bundle, and $T^*\mathscr{U} = \mathscr{T}^0{}_1(\mathscr{U})$ is the dual thereto. Any (k,l)-tensor in any $(3+1)$-dimensional (Lorentzian) spacetime $\mathscr{M}$ is equivalently denoted either using the (boldface) index-free notation $\boldsymbol{A} \in \mathscr{T}^k{}_l(\mathscr{M})$ following the practice of, e.g., Refs. [50,66], or the abstract index notation $A^{a_1\cdots a_k}{}_{b_1\cdots b_l} \in \mathscr{T}^k{}_l(\mathscr{M})$ following that of, e.g., Ref. [67]; that is, depending upon convenience, we equivalently write

$$\boldsymbol{A} = A^{a_1\cdots a_k}{}_{b_1\cdots b_l} \in \mathscr{T}^k{}_l(\mathscr{M}), \tag{12}$$

with latin letters from the beginning of the alphabet $(a,b,c,\ldots)$ being used for spacetime indices (0,1,2,3). The components of $\boldsymbol{A}$ in a particular choice of coordinates $\{x^\alpha\}_{\alpha=0}^3$ are denoted by $A^{\alpha_1\cdots\alpha_k}{}_{\beta_1\cdots\beta_l}$, using greek (rather than latin) letters from the beginning of the alphabet $(\alpha,\beta,\gamma,\ldots)$. Spatial indices on an appropriately defined (three-dimensional Riemannian spacelike) constant time slice of $\mathscr{M}$ are denoted using latin letters from the middle third of the alphabet in roman font: in lowercase $(i,j,k,\ldots)$ if they are abstract and in uppercase $(I,J,K,\ldots)$ if a particular choice of coordinates $\{x^I\}_{I=1}^3$ has been made.

For any n-dimensional manifold $(\mathscr{U},\boldsymbol{g}_{\mathscr{U}},\boldsymbol{\nabla}_{\mathscr{U}})$ with metric $\boldsymbol{g}_{\mathscr{U}}$ and compatible derivative operator $\boldsymbol{\nabla}_{\mathscr{U}}$, we denote its natural volume form by

$$\boldsymbol{\epsilon}_{\mathscr{U}} = \sqrt{|\det(\boldsymbol{g}_{\mathscr{U}})|}\,\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n. \tag{13}$$

Let $\mathscr{S} \simeq \mathbb{S}^2$ be any (Riemannian) closed two-surface that is topologically a two-sphere. Latin letters from the middle third of the alphabet in fraktur font $(\mathfrak{i},\mathfrak{j},\mathfrak{k},\ldots)$ are reserved for indices of tensors in $\mathscr{T}^k{}_l(\mathscr{S})$. In particular, for $\mathbb{S}^2$ itself, $\mathfrak{G}_{\mathfrak{ij}}$ is the metric, $\mathfrak{D}_{\mathfrak{i}}$ is the associated derivative operator, and $\epsilon^{\mathbb{S}^2}_{\mathfrak{ij}}$ is the volume form; in standard spherical coordinates $\{\theta,\phi\}$, the latter is simply given by

$$\boldsymbol{\epsilon}_{\mathbb{S}^2} = \sin\theta\,\mathrm{d}\theta \wedge \mathrm{d}\phi. \tag{14}$$

Contractions are indicated in the usual way in the abstract index notation: e.g., U^aV_a is the contraction of $\boldsymbol{U}$ and $\boldsymbol{V}$. Equivalently, when applicable, we may simply use the "dot product" in the index-free notation, e.g., $U^aV_a = \boldsymbol{U}\cdot\boldsymbol{V}$, $A_{ab}B^{ab} = \boldsymbol{A}:\boldsymbol{\mathscr{B}}$, etc. We must keep in mind that such contractions are to be performed using the metric of the space on which the relevant tensors are defined. Additionally, often we find it convenient to denote the component (projection) of a tensor in a certain direction by simply replacing its pertinent abstract index therewith: e.g., we equivalently write $U^aV_b = \boldsymbol{U}\cdot\boldsymbol{V} = U_{\boldsymbol{V}} = V_{\boldsymbol{U}}$, $A_{ab}U^a = A_{\boldsymbol{U}b}$, $A_{ab}U^aV^b = A_{\boldsymbol{UV}}$, etc. For any (0,2)-tensor A_{ab} that is not a metric, we indicate its trace (in nonboldface) as $A = A_a{}^a = \mathrm{tr}(\boldsymbol{A})$. (If it is a metric, this notation is reserved, as usual, for its determinant rather than its trace.)

Finally, let $\mathscr{U}$ and $\mathscr{V}$ be any two diffeomorphic manifolds and let $f:\mathscr{U}\to\mathscr{V}$ be a map between them. This naturally defines a map between tensors on the two manifolds, which we denote by $f_*:\mathscr{T}^k{}_l(\mathscr{U}) \to \mathscr{T}^k{}_l(\mathscr{V})$ and its inverse $(f^{-1})_* = f^*:\mathscr{T}^k{}_l(\mathscr{V}) \to \mathscr{T}^k{}_l(\mathscr{U})$ if it exists. We generically refer to any map of this sort as a *tensor transport* [68]. It is simply the generalization to arbitrary tensors of the pushforward $f_*:T\mathscr{U}\to T\mathscr{V}$ and pullback $f^*:T^*\mathscr{V}\to T^*\mathscr{U}$, the action of which is defined in the standard way—see, e.g., Appendix C of Ref. [67]. (Note that our convention for sub-/superscripting the star is the generally more common one used in geometry [68,69]; it is sometimes opposite to and sometimes congruous with that used in the physics literature, e.g., Refs. [67] and [70], respectively).

II. SETUP: QUASILOCAL CONSERVATION LAWS

Let $(\mathscr{M},\boldsymbol{g},\boldsymbol{\nabla})$ be any $(3+1)$-dimensional spacetime such that, given any matter stress-energy-momentum tensor T_{ab}, the Einstein equation,

$$\boldsymbol{G} = \kappa \boldsymbol{T} \text{ in } \mathscr{M}, \tag{15}$$

holds. In what follows, we introduce the concept of quasilocal frames [51,55–61] and describe the basic steps for their construction, as well as the energy and momentum conservation laws associated therewith. In Sec. II A, we offer an heuristic idea of quasilocal frames before proceeding in Sec. II B to present the full mathematical construction. Then, in Sec. II C, we motivate and discuss the quasilocal stress-energy-momentum tensor used in this work, that is, the Brown-York tensor. Finally, in Sec. II D, we review the formulation of quasilocal conservation laws using these ingredients.

A. Quasilocal frames: Heuristic idea

Before we enter into the technical details, we would like to offer a heuristic picture and motivation for defining the concept of quasilocal frames.

We would like to show how the GSF arises from general-relativistic conservation laws. For this, we require first the embedding into our spacetime $\mathscr{M}$ of a worldtube boundary $\mathscr{B} \simeq \mathbb{R} \times \mathbb{S}^2$. The worldtube interior contains the system of which the dynamics we are interested in describing. In principle, such a $\mathscr{B}$ can be completely specified by choosing an appropriate *radial function* $r(x)$ on $\mathscr{M}$ and setting it equal to a non-negative constant [such that the $r(x) = \text{const} > 0$ Lorentzian slices of $\mathscr{M}$ have topology $\mathbb{R} \times \mathbb{S}^2$]. This would be analogous to defining a (Riemannian, with topology $\mathbb{R}^3$) Cauchy surface by the constancy of a time function $t(x)$ on $\mathscr{M}$.

However, this does not quite suffice. As we have briefly argued in the Introduction (and will shortly elaborate upon in greater technicality), the conservation laws appropriate to GR ought to be quasilocal in form, that is, involving stress-energy-momentum as boundary (not volume) densities. One may readily assume that the latter are defined by a quasilocal stress-energy-momentum tensor living on $\mathscr{B}$, which we denote—for the moment, generally—by τ_{ab}. (Later, we give an explicit definition, namely, that of the Brown-York tensor, for $\boldsymbol{\tau}$.)

To construct conservation laws, then, one would need to project this $\boldsymbol{\tau}$ into directions on $\mathscr{B}$, giving quantities such as energy or momenta, and then to consider their flux through a portion of $\mathscr{B}$ (an interval of time along the worldtube boundary). But in this case, we have to make clear what is meant by the energy (time-time) and momenta (time-space) components of $\boldsymbol{\tau}$ within $\mathscr{B}$, the changes in which we are interested in studying. For this reason, additional constructions are required.

In particular, what we need is a *congruence* of observers with respect to which projections of $\boldsymbol{\tau}$ yield stress-energy-momentum quantities. Since $\boldsymbol{\tau}$ is only defined on $\mathscr{B}$, this therefore needs to be a two-parameter family of (timelike) worldlines, the union of which is $\mathscr{B}$ itself. This is analogous to how the integral curves of a time flow vector field (as in canonical GR) altogether constitute ("fill up") the entire spacetime $\mathscr{M}$, except that there we are dealing with a three- (rather than two-)parameter family of timelike worldlines.

We refer to any set of observers, the worldlines of which form a two-parameter family constituting $\mathscr{B} \simeq \mathbb{R} \times \mathbb{S}^2$, as *quasilocal* observers. A specification of such a two-parameter family, equivalent to specifying the unit four-velocity $u^a \in T\mathscr{B}$ of these observers (the integral curves of which "trace out" $\mathscr{B}$), is what is meant by a quasilocal frame.

With this, we can now meaningfully talk about projections of $\boldsymbol{\tau}$ into directions on $\mathscr{B}$ as stress-energy-momentum quantities. For example, $\tau_{\boldsymbol{uu}}$ may appear immediately suggestible as a definition for the (boundary) energy density. Indeed, later, we take precisely this definition, and we will furthermore see how momenta (the basis of the GSF problem) can be defined as well.

B. Quasilocal frames: Mathematical construction

Concordant with our discussion in the previous subsection, a quasilocal frame (see Fig. 5 for a graphical illustration of the construction) is defined as a two-parameter family of timelike worldlines constituting the worldtube boundary (topologically $\mathbb{R} \times \mathbb{S}^2$) of the history

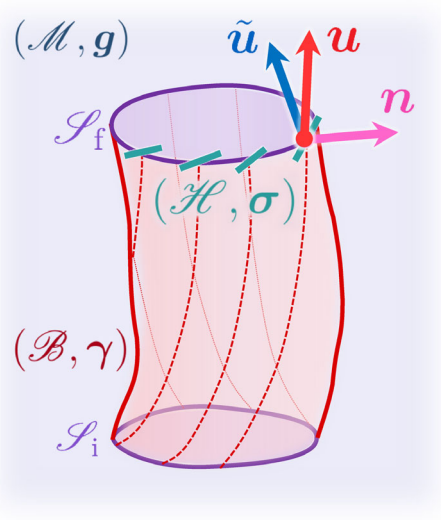

FIG. 5. A portion of a quasilocal frame $(\mathscr{B}; \boldsymbol{u})$ in a spacetime $\mathscr{M}$, bounded by constant t two-surfaces $\mathscr{S}_{\mathrm{i}}$ and $\mathscr{S}_{\mathrm{f}}$. In particular, $\mathscr{B} \simeq \mathbb{R} \times \mathbb{S}^2$ is the union of all integral curves (two-parameter family of timelike worldlines), depicted in the figure as dotted red lines, of the vector field $\boldsymbol{u} \in T\mathscr{B}$ which represents the unit four-velocity of quasilocal observers making up the congruence. The unit normal to $\mathscr{B}$ (in $\mathscr{M}$) is $\boldsymbol{n}$, and the normal to each constant t slice $\mathscr{I}$ of $\mathscr{B}$ is $\tilde{\boldsymbol{u}}$ (not necessarily coincidental with $\boldsymbol{u}$). Finally, $\mathscr{H}$ (with induced metric $\boldsymbol{\sigma}$) is the two-dimensional subspace of $T\mathscr{B}$ consisting of the spatial vectors orthogonal to $\boldsymbol{u}$. Note that unlike $\mathscr{I}$, $\mathscr{H}$ need not be integrable (indicated in the figure by the failure of $\mathscr{H}$ to make a closed two-surface).

of a finite (closed) spatial three-volume in $\mathscr{M}$. Let u^a denote the timelike unit vector field tangent to these worldlines. Such a congruence constitutes a submanifold of $\mathscr{M}$ that we call $\mathscr{B} \simeq \mathbb{R} \times \mathbb{S}^2$. Let n^a be the outward-pointing unit vector field normal to $\mathscr{B}$; note that $\boldsymbol{n}$ is uniquely fixed once $\mathscr{B}$ is specified. There is thus a Lorentzian metric $\boldsymbol{\gamma}$ [of signature $(-,+,+)$] induced on $\mathscr{B}$, the components of which are given by

$$\gamma_{ab} = g_{ab} - n_a n_b. \tag{16}$$

We denote the induced derivative operator compatible therewith by $\boldsymbol{\mathcal{D}}$. To indicate that a topologically $\mathbb{R} \times \mathbb{S}^2$ submanifold $(\mathscr{B}, \boldsymbol{\gamma}, \boldsymbol{\mathcal{D}})$ of $\mathscr{M}$ is a quasilocal frame (that is to say, defined as a particular congruence with four-velocity $\boldsymbol{u}$ as detailed above, and not just as an embedded submanifold) in $\mathscr{M}$, we write $(\mathscr{B}, \boldsymbol{\gamma}, \boldsymbol{\mathcal{D}}; \boldsymbol{u})$ or simply $(\mathscr{B}; \boldsymbol{u})$.

Let $\mathscr{H}$ be the two-dimensional subspace of $T\mathscr{B}$ consisting of the "spatial" vectors orthogonal to $\boldsymbol{u}$. Let $\boldsymbol{\sigma}$ denote the two-dimensional (spatial) Riemannian metric [of signature $(+,+)$] that projects tensor indices into $\mathscr{H}$, and is induced on $\mathscr{B}$ by the choice of $\boldsymbol{u}$ (and thus also $\boldsymbol{n}$), given by

$$\sigma_{ab} = \gamma_{ab} + u_a u_b = g_{ab} - n_a n_b + u_a u_b. \tag{17}$$

The induced derivative operator compatible with $\boldsymbol{\sigma}$ is denoted by $\boldsymbol{D}$. Let $\{x^{\mathfrak{i}}\}_{\mathfrak{i}=1}^2$ (written using fraktur indices from the middle third of the latin alphabet) be spatial coordinates on $\mathscr{B}$ that label the worldlines of the observers, and let t be a time coordinate on $\mathscr{B}$ such that surfaces of constant t, to which there exists a unit normal vector that we denote by $\tilde{u}^a \in T\mathscr{B}$, foliate $\mathscr{B}$ by closed spatial two-surfaces $\mathscr{S}$ (with topology $\mathbb{S}^2$). Letting N denote the lapse function of $\boldsymbol{g}$, we have $\boldsymbol{u} = N^{-1}\partial/\partial t$.

Note that in general, $\mathscr{H}$ need not coincide with the constant time slices $\mathscr{S}$. Equivalently, $\boldsymbol{u}$ need not coincide with $\tilde{\boldsymbol{u}}$. In general, there will be a shift between them, such that

$$\tilde{\boldsymbol{u}} = \tilde{\gamma}(\boldsymbol{u} + \boldsymbol{v}), \tag{18}$$

where v^a represents the spatial two-velocity of fiducial observers that are at rest with respect to $\mathscr{S}$ as measured by our congruence of quasilocal observers (the four-velocity of which is $\boldsymbol{u}$), and $\tilde{\gamma} = 1/\sqrt{1 - \boldsymbol{v}\cdot\boldsymbol{v}}$ is the Lorentz factor.

The specification of a quasilocal frame is thus equivalent to making a particular choice of a two-parameter family of timelike worldlines comprising $\mathscr{B}$. There are, *a priori*, three degrees of freedom (DoF's) available to us for doing this. Heuristically, these can be regarded as corresponding to the three DoF's in choosing the direction of $\boldsymbol{u}$—from which $\boldsymbol{n}$ and all induced quantities are then computable. (Note that $\boldsymbol{u}$ has four components, but one of the four is fixed by the normalization requirement $\boldsymbol{u}\cdot\boldsymbol{u} = -1$, leaving three independent direction DoF's.) Equivalently, we are in principle free to pick any three geometrical conditions (along the congruence) to fix a quasilocal frame. In practice, usually it is physically more natural, as well as mathematically easier, to work with geometric quantities other than $\boldsymbol{u}$ itself to achieve this.

Yet, it is worth remarking that simply writing down three desired equations (or conditions) to be satisfied by geometrical quantities on $\mathscr{B}$ does not itself guarantee that, in general, a submanifold $(\mathscr{B}, \boldsymbol{\gamma}, \boldsymbol{D})$ obeying those three particular equations will always exist—and, if it does, that it will be the unique such submanifold—in an arbitrary $(\mathscr{M}, \boldsymbol{g}, \boldsymbol{\nabla})$. Nevertheless, one choice of quasilocal frame that is known to always exist (a claim we will qualify more carefully in a moment) is that where the two-metric $\boldsymbol{\sigma}$ on $\mathscr{H}$ is "rigid" (or "time independent")—these are called *rigid quasilocal frames*.

Most of the past work on quasilocal frames has in fact been done in the rigid case [51,55–59]. We know, however, that other quasilocal frame choices are also possible, such as *geoids*—dubbed geoid quasilocal frames [60,61]; these are the general-relativistic generalization of "constant gravitational potential" surfaces in Newtonian gravity. Regardless, the quasilocal frame choice that we will mainly consider in this paper is the rigid one (and we will be clear when this choice is explicitly enacted).

Intuitively, the reason for this preference is that imposing in this way the condition of "spatial rigidity" on $(\mathscr{B}; \boldsymbol{u})$—a two-dimensional (boundary) rigidity requirement, which, unlike three-dimensional rigidity, is permissible in GR—eliminates from the description of the system any effects arising simply from the motion of the quasilocal observers relative to each other. Thus, the physics of what is going on inside the system (i.e., the worldtube interior) is essentially all that affects its dynamics.

Technically, there is a further reason: a proof of the existence of solutions—i.e., the existence of a submanifold $\mathscr{B} \simeq \mathbb{R} \times \mathbb{S}^2$ in $\mathscr{M}$ that is also a quasilocal frame $(\mathscr{B}; \boldsymbol{u})$—for any spacetime $(\mathscr{M}, \boldsymbol{g}, \boldsymbol{\nabla})$ has up to now only been fully carried out for rigid quasilocal frames.[9] While, as we have commented, other quasilocal frame choices may be generally possible in principle (and may be shown to be possible to construct, case-by-case, in specific spacetimes—as we have done, e.g., with geoid quasilocal frames [60,61]), they are as yet not rigorously guaranteed to exist in arbitrary spacetimes.

The quasilocal rigidity conditions can be stated in a number of ways. Most generally, defining

$$\theta_{ab} = \sigma_{ac}\sigma_{bd}\nabla^c u^d \tag{19}$$

to be the *strain rate tensor* of the congruence, they amount to the requirement of vanishing expansion $\theta = \mathrm{tr}(\boldsymbol{\theta})$ and shear $\theta_{\langle ab\rangle} = \theta_{(ab)} - \frac{1}{2}\theta\sigma_{ab}$, i.e.,

[9]The idea of the proof is to explicitly construct the solutions order by order in an expansion in the areal radius around an arbitrary worldline in an arbitrary spacetime [57].

$$\theta = 0 = \theta_{\langle ab\rangle} \Leftrightarrow 0 = \theta_{(ab)}. \tag{20}$$

In the adapted coordinates, these three conditions are expressible as the vanishing of the time derivative of the two-metric on $\mathscr{H}$, i.e., $0 = \partial_t \boldsymbol{\sigma}$. Both of these two equivalent mathematical conditions, $\theta_{(ab)} = 0 = \partial_t \boldsymbol{\sigma}$, capture physically the meaning of the quasilocal observers moving rigidly with respect to each other (i.e., the "radar-ranging" distances between them does not change in time).

C. Quasilocal stress-energy-momentum tensor

Before we consider the formulation of conservation laws with the use of quasilocal frames (from which our analysis of the GSF will eventually emerge), we wish to address in a bit more detail an even more fundamental question: what are conservation laws in GR actually supposed to be about? At the most basic level, they should express changes (over time) in some appropriately defined notion of energy-momentum. As we are interested in gravitational systems (and specifically, those driven by the effect of the GSF), this energy-momentum must include that of the gravitational field, in addition to that of any matter fields if present.

Hence, we may assert from the outset that it does not make much sense in GR to seek conservation laws based solely on the matter stress-energy-momentum tensor $\boldsymbol{T}$, such as Eq. (8). It is evident that these would, by construction, account for matter only—leaving out gravitational effects in general (which could exist in the complete absence of matter, e.g., gravitational waves), and thus the GSF in particular. What is more, such conservation laws are logically inconsistent from a general-relativistic point of view: a nonvanishing $\boldsymbol{T}$ implies a nontrivial gravitational field (through the Einstein equation) and thus a necessity of taking into account that field along with the matter one(s) for a proper accounting of energy-momentum transfer. A further technical problem is also that the formulation of conservation laws of this sort is typically predicated upon the existence of Killing vector fields or other types of symmetry generators in $\mathscr{M}$, which one does not have in general—and which do not exist in spacetimes pertinent for the GSF problem in particular.

We are therefore led to ask how we can meaningfully define a total—*gravity plus matter*—stress-energy-momentum tensor in GR. It turns out that the precise answer to this question, while certainly not intractable, is unfortunately also not unique—or at least, it lacks a clear consensus among relativists, even today. See, e.g., Refs. [52,53] for reviews of the variety of proposals that have been put forward toward addressing this question. Nonetheless, for reasons already touched upon and to be elaborated presently, what is clear and generally accepted is that such a tensor cannot be local in nature (as $\boldsymbol{T}$ is) and for this reason is referred to as quasilocal.

Let τ_{ab} denote this quasilocal, total (matter plus gravity) stress-energy-momentum tensor that we eventually seek to use for our conservation laws. It has long been understood [50] that whatever the notion of gravitational energy-momentum (defined by $\boldsymbol{\tau}$) might mean, it is not something localizable; in other words, there is no way of meaningfully defining an "energy-momentum volume density" for gravity. This is, ultimately, due to the equivalence principle: locally, one can always find a reference frame in which all local gravitational fields (the connection coefficients), and thus any notion of energy-momentum volume density associated therewith, disappear. The remedy is to make $\boldsymbol{\tau}$ quasilocal, meaning that, rather than volume densities, it should define surface densities (of energy, momentum, etc.)—a type of construction which *is* mathematically realizable and physically sensible in general.

The specific choice we make for how to define this total (matter plus gravity), quasilocal energy-momentum tensor $\boldsymbol{\tau}$ is the so-called Brown-York tensor, first put forward by the authors in Ref. [54]; see also Ref. [71] for a detailed review. This proposal was based originally upon a Hamilton-Jacobi analysis; here, we will offer a simpler argument for its definition, sketched out initially in Ref. [51].

Consider the standard gravitational action $S_{\rm G}$ for a spacetime volume $\mathscr{V} \subset \mathscr{M}$ such that $\partial\mathscr{V} = \mathscr{B} \simeq \mathbb{R} \times \mathbb{S}^2$ is a worldtube boundary as in the previous subsection (possibly constituting a quasilocal frame, but not necessarily). This action is given by the sum of two terms, a bulk and a boundary term, respectively:

$$S_{\rm G}[\boldsymbol{g}] = S_{\rm EH}[\boldsymbol{g}] + S_{\rm GHY}[\boldsymbol{\gamma}, \boldsymbol{n}]. \tag{21}$$

In particular, the first is the Einstein-Hilbert bulk term,

$$S_{\rm EH}[\boldsymbol{g}] = \frac{1}{2\kappa}\int_{\mathscr{V}} \boldsymbol{\epsilon}_{\mathscr{M}} R, \tag{22}$$

and the second is the Gibbons-Hawking-York boundary term [72,73],

$$S_{\rm GHY}[\boldsymbol{\gamma}, \boldsymbol{n}] = -\frac{1}{\kappa}\int_{\partial\mathscr{V}} \boldsymbol{\epsilon}_{\mathscr{B}} \Theta. \tag{23}$$

Here, $\boldsymbol{\epsilon}_{\mathscr{M}} = \mathrm{d}^4x\sqrt{-g}$ is the volume form on $\mathscr{M}$ with $g = \det(\boldsymbol{g})$, $\boldsymbol{\epsilon}_{\mathscr{B}} = \mathrm{d}^3x\sqrt{-\gamma}$ is the volume form on $\mathscr{B}$ with $\gamma = \det(\boldsymbol{\gamma})$, and $\Theta = \mathrm{tr}(\boldsymbol{\Theta})$ is the trace of the extrinsic curvature $\Theta_{ab} = \gamma_{ac}\nabla^c n_b$ of $\mathscr{B}$ in $\mathscr{M}$. Additionally, the matter action $S_{\rm M}$ for any set of matter fields Ψ described by a Lagrangian $L_{\rm M}$ is

$$S_{\rm M}[\Psi] = \int_{\mathscr{V}} \boldsymbol{\epsilon}_{\mathscr{M}} L_{\rm M}[\Psi]. \tag{24}$$

The definition of the total (quasilocal) stress-energy-momentum tensor $\boldsymbol{\tau}$ for gravity plus matter can be obtained effectively in the same way as that of the (local) stress-energy-momentum tensor $\boldsymbol{T}$ for matter alone—from the

total action in Eq. (21) rather than just, respectively, the matter action in Eq. (24). In particular, $\boldsymbol{T}$ is defined by computing the variation δ (with respect to the spacetime metric) of the matter action:

$$\delta S_{\mathrm{M}}[\Psi] = -\frac{1}{2}\int_{\mathscr{V}} \boldsymbol{\epsilon}_{\mathscr{M}} T_{ab}\delta g^{ab}. \tag{25}$$

In other words, one defines the matter stress-energy-momentum tensor as the functional derivative,

$$T_{ab} = -\frac{2}{\sqrt{-g}}\frac{\delta S_{\mathrm{M}}}{\delta g^{ab}}. \tag{26}$$

The definition of the Brown-York tensor follows completely analogously, except that now gravity is also included. That is, for the total action of gravity (minimally) coupled to matter,

$$S_{\mathrm{G+M}}[\boldsymbol{g},\Psi] = S_{\mathrm{G}}[\boldsymbol{g}] + S_{\mathrm{M}}[\Psi], \tag{27}$$

we have that the metric variation is

$$\delta S_{\mathrm{G+M}}[\boldsymbol{g},\Psi] = \frac{1}{2}\left\{\int_{\mathscr{V}} \boldsymbol{\epsilon}_{\mathscr{M}}\left(\frac{1}{\kappa}G_{ab} - T_{ab}\right)\delta g^{ab} - \int_{\partial\mathscr{V}} \boldsymbol{\epsilon}_{\mathscr{B}}\left(-\frac{1}{\kappa}\Pi_{ab}\right)\delta\gamma^{ab}\right\} \tag{28}$$

$$= -\frac{1}{2}\int_{\partial\mathscr{V}} \boldsymbol{\epsilon}_{\mathscr{B}}\tau_{ab}\delta\gamma^{ab}. \tag{29}$$

In the equality of Eq. (28), $\boldsymbol{\Pi}$ is the canonical momentum of $(\mathscr{B},\boldsymbol{\gamma},\boldsymbol{\mathcal{D}})$, given by $\boldsymbol{\Pi} = \boldsymbol{\Theta} - \Theta\boldsymbol{\gamma}$. It follows from direct computation using Eqs. (21), (24), and (25); for a review of this derivation carefully accounting for the boundary term see, e.g., Chapter 12 of Ref. [74]. In the equality of Eq. (29), the Einstein equation $\boldsymbol{G} = \kappa\boldsymbol{T}$ has been invoked (in other words, we impose the Einstein equation to be satisfied in the bulk), thus leading to the vanishing of the bulk term; meanwhile in the boundary term, a gravity plus matter stress-energy-momentum tensor $\boldsymbol{\tau}$ (the Brown-York tensor) has been defined in direct analogy with the definition of the matter energy-momentum tensor $\boldsymbol{T}$ in Eq. (25). Hence, just as Eq. (25) implies Eq. (26), Eq. (29) implies

$$\boldsymbol{\tau} = -\frac{1}{\kappa}\boldsymbol{\Pi}. \tag{30}$$

Henceforth, $\boldsymbol{\tau}$ refers strictly to this (Brown-York) quasilocal stress-energy-momentum tensor of Eq. (30), and not to any other definition.

It is useful to decompose $\boldsymbol{\tau}$ in a similar way as is ordinarily done with $\boldsymbol{T}$, so we define

$$\mathcal{E} = u^a u^b \tau_{ab}, \tag{31}$$

$$\mathcal{P}^a = -\sigma^{ab}u^c\tau_{bc}, \tag{32}$$

$$\mathcal{S}^{ab} = -\sigma^{ac}\sigma^{bd}\tau_{cd}, \tag{33}$$

as the quasilocal energy, momentum, and stress, respectively, with units of energy per unit area, momentum per unit area, and force per unit length. Equivalently,

$$\tau^{ab} = u^a u^b \mathcal{E} + 2u^{(a}\mathcal{P}^{b)} - \mathcal{S}^{ab}. \tag{34}$$

D. Conservation laws

The construction of general conservation laws from $\boldsymbol{\tau}$ was first achieved in Refs. [51,56], and proceeds along the following lines. Let $\boldsymbol{\psi} \in T\mathscr{B}$ be an arbitrary vector field in $\mathscr{B}$. We begin by considering a projection of $\boldsymbol{\Pi}$ in the direction of $\boldsymbol{\psi}$ (in one index), i.e., $\Pi^{ab}\psi_b$, and computing its divergence in $\mathscr{B}$. By using the Leibnitz rule, we simply have

$$\mathcal{D}_a(\Pi^{ab}\psi_b) = (\mathcal{D}_a\Pi^{ab})\psi_b + \Pi^{ab}(\mathcal{D}_a\psi_b). \tag{35}$$

Next, we integrate this equation over a portion $\Delta\mathscr{B}$ of $\mathscr{B}$ bounded by initial and final constant t surfaces $\mathscr{I}_{\mathrm{i}}$ and $\mathscr{I}_{\mathrm{f}}$, as depicted in Fig. 5. On the resulting lhs, we apply Stokes's theorem, and on the first term on the rhs, we use the Gauss-Codazzi identity: $\mathcal{D}_a\Pi^{ab} = n_a\gamma^b_c G^{ac}$. Thus, using the notation for tensor projections in certain directions introduced in Sec. I E for ease of readability (e.g., $G_{ab}n^a\psi^b = G_{\boldsymbol{n}\boldsymbol{\psi}}$ and similarly for other contractions), we obtain

$$\int_{\mathscr{I}_{\mathrm{f}}-\mathscr{I}_{\mathrm{i}}} \boldsymbol{\epsilon}_{\mathscr{I}}\Pi_{\tilde{\boldsymbol{u}}\boldsymbol{\psi}} = -\int_{\Delta\mathscr{B}} \boldsymbol{\epsilon}_{\mathscr{B}}(G_{\boldsymbol{n}\boldsymbol{\psi}} + \Pi^{ab}\mathcal{D}_a\psi_b), \tag{36}$$

where $\boldsymbol{\epsilon}_{\mathscr{I}}$ denotes the volume form on the constant time closed two-surfaces $\mathscr{I}$, and we have used the notation $\int_{\mathscr{I}_{\mathrm{f}}-\mathscr{I}_{\mathrm{i}}}(\cdot) = \int_{\mathscr{I}_{\mathrm{f}}}(\cdot) - \int_{\mathscr{I}_{\mathrm{i}}}(\cdot)$. We also remind the reader that $\tilde{\boldsymbol{u}}$ represents the unit normal to each constant time closed two-surface, which in general need not coincide with the quasilocal observers' four-velocity $\boldsymbol{u}$ but is related to it by a Lorentz transformation, Eq. (18); see also Fig. 5.

We stress that, so far, Eq. (36) is a purely geometrical identity, completely general for any Lorentzian manifold $\mathscr{M}$; in other words, thus far, we have said nothing about physics.

Now, to give this identity physical meaning, we invoke the definition of the Brown-York tensor in Eq. (30) (giving the boundary extrinsic geometry its meaning as stress-energy-momentum) as well as the Einstein equation [Eq. (15)], giving the spacetime curvature its meaning as the gravitational field. With these, Eq. (36) turns into

$$\int_{\mathscr{I}_{\rm f}-\mathscr{I}_{\rm i}} \boldsymbol{\epsilon}_{\mathscr{I}} \tilde{\gamma}(\tau_{\boldsymbol{u}\boldsymbol{\psi}} + \tau_{\boldsymbol{v}\boldsymbol{\psi}}) = \int_{\Delta\mathscr{B}} \boldsymbol{\epsilon}_{\mathscr{B}} (T_{\boldsymbol{n}\boldsymbol{\psi}} - \tau^{ab}\mathcal{D}_{(a}\psi_{b)}). \tag{37}$$

On the lhs, we have inserted the relation $\tilde{\boldsymbol{u}} = \tilde{\gamma}(\boldsymbol{u} + \boldsymbol{v})$, with v^a representing the spatial two-velocity of fiducial observers that are at rest with respect to $\mathscr{I}$ (the hypersurface-orthogonal four-velocity of which is $\tilde{\boldsymbol{u}}$) as measured by our congruence of quasilocal observers (the four-velocity of which is $\boldsymbol{u}$), and $\tilde{\gamma} = 1/\sqrt{1 - \boldsymbol{v}\cdot\boldsymbol{v}}$ is the Lorentz factor.

Observe that Eq. (37) expresses the change of some component of the quasilocal stress-energy-momentum tensor integrated over two different $t = \text{const}$ closed two-surfaces $\mathscr{I}$ as a flux through the worldtube boundary $\Delta\mathscr{B}$ between them. The identification of the different components of $\boldsymbol{\tau}$ as the various components of the total energy-momentum of the system thus leads to the understanding of Eq. (37) as a general conservation law for the system contained inside of $\Delta\mathscr{B}$. Thus, depending on our particular choice of $\boldsymbol{\psi} \in T\mathscr{B}$, Eq. (37) will represent a conservation law for the total energy, momentum, or angular momentum of this system [51].

Let us now assume that $(\mathscr{B}; \boldsymbol{u})$ is a rigid quasilocal frame. If we choose $\boldsymbol{\psi} = \boldsymbol{u}$, then Eq. (37) becomes the energy conservation law,

$$\int_{\mathscr{I}_{\rm f}-\mathscr{I}_{\rm i}} \boldsymbol{\epsilon}_{\mathscr{I}} \tilde{\gamma}(\mathcal{E} - \mathcal{P}_{\boldsymbol{v}}) = \int_{\Delta\mathscr{B}} \boldsymbol{\epsilon}_{\mathscr{B}} (T_{\boldsymbol{n}\boldsymbol{u}} - \boldsymbol{\alpha}\cdot\boldsymbol{\mathcal{P}}), \tag{38}$$

where $\alpha^a = \sigma^{ab}a_b$ is the $\mathscr{H}$ projection of the acceleration of the quasilocal observers, defined by $a^a = \nabla_{\boldsymbol{u}}u^a$.

Now suppose, on the other hand, that we instead choose $\boldsymbol{\psi} = -\boldsymbol{\phi}$ where $\boldsymbol{\phi} \in \mathscr{H}$ is orthogonal to $\boldsymbol{u}$ (with the minus sign introduced for convenience), and represents a stationary conformal Killing vector field with respect to $\boldsymbol{\sigma}$. This means that $\boldsymbol{\phi}$ is chosen such that it satisfies the conformal Killing equation, $\mathcal{L}_{\boldsymbol{\phi}}\boldsymbol{\sigma} = (\boldsymbol{D}\cdot\boldsymbol{\phi})\boldsymbol{\sigma}$, with $\mathcal{L}$ the Lie derivative and $\boldsymbol{D}$ the derivative on $\mathscr{H}$ (compatible with $\boldsymbol{\sigma}$). A set of six such conformal Killing vectors always exists: three for translations and three for rotations, respectively generating the action of boosts and rotations of the Lorentz group on the two-sphere [51]. The idea, then, is that the contraction of these vectors with the quasilocal momentum integrated over a constant-time topological two-sphere boundary expresses, respectively, the total linear and angular momentum (in the three ordinary spatial directions each) at that time instant. Thus, Eq. (37) becomes the (respectively, linear and angular) momentum conservation law,

$$\int_{\mathscr{I}_{\rm f}-\mathscr{I}_{\rm i}} \boldsymbol{\epsilon}_{\mathscr{I}} \tilde{\gamma}(\mathcal{P}_{\boldsymbol{\phi}} + \mathcal{S}_{\boldsymbol{v}\boldsymbol{\phi}})$$
$$= -\int_{\Delta\mathscr{B}} \boldsymbol{\epsilon}_{\mathscr{B}} (T_{\boldsymbol{n}\boldsymbol{\phi}} + \mathcal{E}\alpha_{\boldsymbol{\phi}} + 2\nu\epsilon_{ab}\mathcal{P}^a\phi^b + \mathrm{P}\boldsymbol{D}\cdot\boldsymbol{\phi}), \tag{39}$$

where $\nu = \frac{1}{2}\epsilon^{ab}_{\mathscr{H}}\mathcal{D}_a u_b$ is the twist of the congruence (with $\epsilon^{\mathscr{H}}_{ab} = \epsilon^{\mathscr{M}}_{abcd}u^c n^d$ the induced volume form on $\mathscr{H}$) and $\mathrm{P} = \frac{1}{2}\boldsymbol{\sigma}:\boldsymbol{\mathcal{S}}$ is the quasilocal pressure (force per unit length) between the worldlines of $\mathscr{B}$. We remark that the latter can be shown to satisfy the very useful general identity (which we will expediently invoke in our later calculations):

$$\mathcal{E} - 2\mathrm{P} = \frac{2}{\kappa}a_{\boldsymbol{n}}. \tag{40}$$

An analysis of the gravitational self-force problem should consider the conservation law in Eq. (39) for *linear momentum*. Thus, we will use the fact, described in greater detail in Appendix A, that the conformal Killing vector $\boldsymbol{\phi} \in \mathscr{H}$ for linear momentum admits a multipole decomposition of the following form:

$$\phi^{\mathfrak{i}} = \frac{1}{r}D^{\mathfrak{i}}(\Phi^I r_I + \Phi^{IJ}r_I r_J + \cdots) \tag{41}$$

$$= \frac{1}{r}(\Phi^I \mathscr{B}^{\mathfrak{i}}_I + 2\Phi^{IJ}\mathscr{B}^{\mathfrak{i}}_I r_J + \cdots), \tag{42}$$

with the dots indicating higher harmonics. Here, r is the area radius of the quasilocal frame (such that $\mathscr{B}$ is a constant r hypersurface in $\mathscr{M}$), r^I denotes the standard direction cosines of a radial unit vector in $\mathbb{R}^3$, and $\mathfrak{B}^{\mathfrak{i}}_I = \partial^{\mathfrak{i}}r_I$ are the boost generators on the two-sphere. See Appendix A for a detailed discussion regarding conformal Killing vectors and the two-sphere. In spherical coordinates $\{\theta, \phi\}$, we have $r^I = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. Thus, Eq. (41) gives us a decomposition of $\boldsymbol{\phi}$ in terms of multipole moments, with the $\ell = 1$ coefficients Φ^I simply representing vectors in $\mathbb{R}^3$ in the direction of which we are considering the conservation law.

III. GENERAL DERIVATION OF THE GRAVITATIONAL SELF-FORCE FROM QUASILOCAL CONSERVATION LAWS

In this section, we will show how the GSF is a general consequence of the momentum conservation law in Eq. (39) for any system which is sufficiently localized. By that, we mean something very simple: taking the $r \to 0$ limit of a quasilocal frame around the moving object which is treated as small, i.e., as a formal perturbation about some background. No further assumptions are for the moment needed. In particular, we do not even need to enter into the precise details of how to specify the perturbation family for this problem; that will be left to the following section, where we will carefully define and work with the family of perturbed spacetimes typically employed for applications of the GSF.

We first review the basic formulation of perturbation theory in GR in Sec. III A. While this material is well known, we find it useful to include it here both for

establishing notation as well as carefully defining the concepts that we need to work with at an adequate level of rigour. Then, in Sec. III B, we show that the first-order perturbation of the momentum conservation law in Eq. (39) always contains the GSF, and that it dominates the dynamics for localized systems.

A. Perturbation theory in GR

Our exposition of perturbation theory in this subsection follows closely the treatment of Ref. [75]. (See also Chapter 7 of Ref. [67] for a simpler treatment of this topic but following the same philosophy.)

Perturbation theory in GR is best made sense of from the point of view of "stacked" manifolds off some known background. To be more precise, let $\lambda \geq 0$ represent our perturbation parameter. It is a purely formal parameter, in the sense that it should be set equal to 1 at the end of any computation and serves only to indicate the order of the perturbation. The idea, then, is to define a one-parameter family of spacetimes $\{(\mathscr{M}_{(\lambda)}, \boldsymbol{g}_{(\lambda)}, \boldsymbol{\nabla}_{(\lambda)})\}_{\lambda \geq 0}$, where $\boldsymbol{\nabla}_{(\lambda)}$ is the connection compatible with the metric $\boldsymbol{g}_{(\lambda)}$ in $\mathscr{M}_{(\lambda)}$, $\forall\, \lambda \geq 0$, such that $(\mathscr{M}_{(0)}, \boldsymbol{g}_{(0)}, \boldsymbol{\nabla}_{(0)}) = (\mathring{\mathscr{M}}, \mathring{\boldsymbol{g}}, \mathring{\boldsymbol{\nabla}})$ is a known, exact spacetime—the background. See Fig. 6 for a visual depiction. For notational convenience, any object with a subscripted "(0)" (from a one-parameter perturbative family) is equivalently written with an overset "$\circ$" instead. For the GSF problem, $\mathring{\mathscr{M}}$ is usually the Schwarzschild-Droste or Kerr spacetime. Then, one should establish a way of smoothly relating the elements of this one-parameter family (between each other) such that calculations on any $\mathscr{M}_{(\lambda)}$ for $\lambda > 0$—which may be, in principle, intractable analytically—can be mapped to calculations on $\mathring{\mathscr{M}}$ in the form of infinite (Taylor) series in λ—which, provided $\mathring{\mathscr{M}}$ is chosen to be a known, exact spacetime, become tractable, order by order, in λ.

Thus, one begins by defining a (five-dimensional, Lorentzian) product manifold,

$$\mathscr{N} = \mathscr{M}_{(\lambda)} \times \mathbb{R}^{\geq}, \tag{43}$$

the natural differentiable structure of which is given simply by the direct product of those on $\mathscr{M}_{(\lambda)}$ and the non-negative real numbers (labeling the perturbation parameter), $\mathbb{R}^{\geq} = \{\lambda \in \mathbb{R} | \lambda \geq 0\}$. For any one-parameter family of (k,l)-tensors $\{\boldsymbol{A}_{(\lambda)}\}_{\lambda \geq 0}$ such that $\boldsymbol{A}_{(\lambda)} \in \mathscr{T}^k{}_l(\mathscr{M}_{(\lambda)})$, $\forall\, \lambda \geq 0$, we define $\boldsymbol{A} \in \mathscr{T}^k{}_l(\mathscr{N})$ by the relation $\mathsf{A}^{\alpha_1 \cdots \alpha_k}{}_{\beta_1 \cdots \beta_l}(p, \lambda) = A_{(\lambda)}{}^{\alpha_1 \cdots \alpha_k}{}_{\beta_1 \cdots \beta_l}(p)$, $\forall\, p \in \mathscr{M}_{(\lambda)}$ and $\forall\, \lambda \geq 0$. Henceforth, any such tensor living on the product manifold will be denoted in serif font—instead of roman font, which remains reserved for tensors living on $(3+1)$-dimensional spacetimes. Furthermore, any spacetime tensor (except for volume forms) or operator written without a sub- or superscripted (λ) lives on $\mathring{\mathscr{M}}$. Conversely, any tensor (except for volume forms) or operator living on $\mathscr{M}_{(\lambda)}$, $\forall\, \lambda > 0$, is indicated via a sub- or (equivalently, if notationally more convenient) superscripted (λ); e.g., $\boldsymbol{A}_{(\lambda)} = \boldsymbol{A}^{(\lambda)} \in \mathscr{T}^k{}_l(\mathscr{M}_{(\lambda)})$ is always tensor in $\mathscr{M}_{(\lambda)}$. The volume form of any (sub)manifold $\mathscr{U}$ is always simply denoted by the standard notation $\boldsymbol{\epsilon}_{\mathscr{U}}$ (and is always understood to live on $\mathscr{U}$), as in Eq. (13).

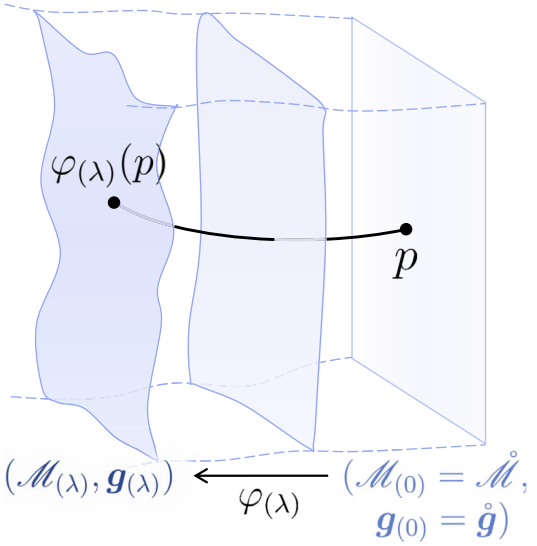

FIG. 6. Representation of a one-parameter family of spacetimes $\{\mathscr{M}_{(\lambda)}\}_{\lambda \geq 0}$ used for perturbation theory. Each of the $\mathscr{M}_{(\lambda)}$ are depicted visually in $(1+1)$ dimensions, as leaves of a (five-dimensional) product manifold $\mathscr{N} = \mathscr{M}_{(\lambda)} \times \mathbb{R}$, with the coordinate $\lambda \geq 0$ representing the perturbative expansion parameter. A choice of a map (or gauge) $\varphi_{(\lambda)} : \mathring{\mathscr{M}} \to \mathscr{M}_{(\lambda)}$ gives us a way of identifying any point $p \in \mathring{\mathscr{M}} = \mathscr{M}_{(0)}$ on the background to one on some perturbed ($\lambda > 0$) spacetime, i.e., $p \mapsto \varphi_{(\lambda)}(p)$.

Let $\Phi^{\mathsf{X}}_{(\lambda)} : \mathscr{N} \to \mathscr{N}$ be a one-parameter group of diffeomorphisms generated by a vector field $\mathsf{X} \in T\mathscr{N}$. (That is to say, the integral curves of X define a flow on $\mathscr{N}$ which connects any two leaves of the product manifold.) For notational convenience, we denote its restriction to maps from the background to a particular perturbed spacetime (identified by a particular value of $\lambda > 0$) as

$$\varphi^{\mathsf{X}}_{(\lambda)} = \Phi^{\mathsf{X}}_{(\lambda)}|_{\mathring{\mathscr{M}}} : \mathring{\mathscr{M}} \to \mathscr{M}_{(\lambda)} \tag{44}$$

$$p \mapsto \varphi^{\mathsf{X}}_{(\lambda)}(p). \tag{45}$$

The choice of X—equivalently, the choice of $\varphi^{\mathsf{X}}_{(\lambda)}$—is not unique; there exists freedom in choosing it, and for this reason, X—equivalently, $\varphi^{\mathsf{X}}_{(\lambda)}$—is referred to as the *perturbative gauge*. We may work with any different gauge choice Y generating a different map $\varphi^{\mathsf{Y}}_{(\lambda)} : \mathring{\mathscr{M}} \to \mathscr{M}_{(\lambda)}$. If we

do not need to render the issue of gauge specification explicit, we may drop the superscript and, instead of $\varphi^{\mathbf{X}}_{(\lambda)}$, simply write $\varphi_{(\lambda)}$.

Consider now the transport under $\varphi^{\mathbf{X}}_{(\lambda)}$ of any tensor $\boldsymbol{A}_{(\lambda)} \in \mathscr{T}^k{}_l(\mathscr{M}_{(\lambda)})$ from a perturbed spacetime to the background manifold. We always denote the transport of any such a tensor by simply dropping the (λ) sub- or superscript and optionally including a superscript to indicate the gauge—that is, $\forall\, \boldsymbol{A}_{(\lambda)} \in \mathscr{T}^k{}_l(\mathscr{M}_{(\lambda)})$,

$$(\varphi^{\mathbf{X}}_{(\lambda)})^* \boldsymbol{A}_{(\lambda)} = \boldsymbol{A}^{\mathbf{X}} = \boldsymbol{A} \in \mathscr{T}^k{}_l(\mathring{\mathscr{M}}), \tag{46}$$

and similarly the transport of $\boldsymbol{\nabla}_{(\lambda)}$ to $\mathring{\mathscr{M}}$ is $\boldsymbol{\nabla}$. We know, moreover, that we can express any such $\boldsymbol{A}$ as a Taylor series around its background value, $\boldsymbol{A}_{(0)} = \mathring{\boldsymbol{A}}$, in Lie derivatives along $\mathbf{X}$ (see Ref. [75]):

$$\boldsymbol{A} = \mathring{\boldsymbol{A}} + \sum_{n=1}^{\infty} \frac{\lambda^n}{n!} \mathcal{L}^n_{\mathbf{X}} \mathbf{A}\big|_{\mathring{\mathscr{M}}} \tag{47}$$

$$= \mathring{\boldsymbol{A}} + \sum_{n=1}^{\infty} \lambda^n \delta^n \boldsymbol{A}, \tag{48}$$

where $\mathcal{L}$ denotes the Lie derivative; in the last equality, we have defined $\delta^n \boldsymbol{A} = (1/n!)(\partial^n_\lambda \boldsymbol{A})|_{\lambda=0}$, and so the (gauge-dependent) first-order perturbation is $\delta^1 \boldsymbol{A} = \delta \boldsymbol{A} = \delta \boldsymbol{A}^{\mathbf{X}}$. Note that the symbol δ^n, $\forall\, n$, can be thought of as an operator $\delta^n = (1/n!)\partial^n_\lambda|_{\lambda=0}$ that acts upon and extracts the $\mathcal{O}(\lambda^n)$ part of any tensor in $\mathring{\mathscr{M}}$. So now, in particular, we have that the background value of $\boldsymbol{g} = (\varphi^{\mathbf{X}}_{(\lambda)})^* \boldsymbol{g}_{(\lambda)}$ is $\mathring{\boldsymbol{g}}$, and we denote its first-order perturbation for convenience and according to convention as $\boldsymbol{h} = \delta \boldsymbol{g}$. Thus, we have

$$\boldsymbol{g} = \mathring{\boldsymbol{g}} + \lambda \boldsymbol{h} + \mathcal{O}(\lambda^2), \tag{49}$$

where we have omitted explicitly specifying the gauge ($\mathbf{X}$) dependence for now.

Let us define one further piece of notation that we shall need to use: let $\mathring{\boldsymbol{\Gamma}}$ and $\boldsymbol{\Gamma} = (\varphi^{X}_{(\lambda)})^* \boldsymbol{\Gamma}_{(\lambda)}$ denote the Christoffel symbols (living on $\mathring{\mathscr{M}}$) associated, respectively, with $\mathring{\boldsymbol{g}}$ and $\boldsymbol{g}$, defined in the usual way (as the connection coefficients between their respective covariant derivatives and the partial derivative). Then, their difference $\boldsymbol{C} = \boldsymbol{\Gamma} - \mathring{\boldsymbol{\Gamma}}$ is the connection coefficient relating $\boldsymbol{\nabla}$ and $\mathring{\boldsymbol{\nabla}}$ on $\mathring{\mathscr{M}}$, which is in fact a tensor. Note that $\mathring{\boldsymbol{C}} = 0$, *i.e.*, $\boldsymbol{C} = \lambda \delta \boldsymbol{C} + \mathcal{O}(\lambda^2)$. In particular, it is given by

$$C^a{}_{bc} = \frac{\lambda}{2} \mathring{g}^{ad} (\mathring{\nabla}_b h_{cd} + \mathring{\nabla}_c h_{bd} - \mathring{\nabla}_d h_{bc}) + \mathcal{O}(\lambda^2). \tag{50}$$

B. Gravitational self-force from the general momentum conservation law

Let $\{(\mathscr{B}_{(\lambda)}; \boldsymbol{u}_{(\lambda)})\}_{\lambda \geq 0}$ be an arbitrary one-parameter family of quasilocal frames (defined as in Sec. II) each of which is embedded, respectively, in the corresponding element of the family of perturbed spacetimes $\{(\mathscr{M}_{(\lambda)}, \boldsymbol{g}_{(\lambda)}, \boldsymbol{\nabla}_{(\lambda)})\}_{\lambda \geq 0}$ described in the previous subsection (Fig. 77). Consider the general geometrical identity (36) in $\mathscr{M}_{(\lambda)}$, $\forall\, \lambda \geq 0$:

$$\int_{\mathscr{T}^{(\lambda)}_{\mathrm{f}} - \mathscr{T}^{(\lambda)}_{\mathrm{i}}} \boldsymbol{\epsilon}_{\mathscr{T}_{(\lambda)}} \Pi^{(\lambda)}_{\tilde{\boldsymbol{u}}_{(\lambda)} \boldsymbol{\psi}_{(\lambda)}} = - \int_{\Delta \mathscr{B}_{(\lambda)}} \boldsymbol{\epsilon}_{\mathscr{B}_{(\lambda)}} (G^{(\lambda)}_{\boldsymbol{n}_{(\lambda)} \boldsymbol{\psi}_{(\lambda)}} + \Pi^{ab}_{(\lambda)} \mathcal{D}^{(\lambda)}_a \psi^{(\lambda)}_b). \tag{51}$$

For $\lambda = 0$, this gives us our conservation laws in the background for $\lambda > 0$, and in a perturbed spacetime for $\lambda > 0$. It is the latter that we are interested in, but since we do not know how to do calculations in $\mathscr{M}_{(\lambda)}$ $\forall\, \lambda > 0$, we have to work with Eq. (51) transported to $\mathring{\mathscr{M}}$. This is easily achieved by using the fact that for any diffeomorphism $f : \mathscr{U} \to \mathscr{V}$ between two (oriented) smooth n-dimensional manifolds $\mathscr{U}$ and $\mathscr{V}$ and any (compactly supported) n-form $\boldsymbol{\omega}$ in $\mathscr{V}$, we have that $\int_{\mathscr{V}} \boldsymbol{\omega} = \int_{\mathscr{U}} f^* \boldsymbol{\omega}$. Applying this to the lhs and rhs of Eq. (51), respectively, we simply get

$$\int_{\varphi^{-1}_{(\lambda)}(\mathscr{T}^{(\lambda)}_{\mathrm{f}}) - \varphi^{-1}_{(\lambda)}(\mathscr{T}^{(\lambda)}_{\mathrm{i}})} (\varphi^*_{(\lambda)} \boldsymbol{\epsilon}_{\mathscr{T}_{(\lambda)}}) \varphi^*_{(\lambda)} \Pi^{(\lambda)}_{\tilde{\boldsymbol{u}}_{(\lambda)} \boldsymbol{\psi}_{(\lambda)}} = \int_{\varphi^{-1}_{(\lambda)}(\Delta \mathscr{B}_{(\lambda)})} (\varphi^*_{(\lambda)} \boldsymbol{\epsilon}_{\mathscr{B}_{(\lambda)}}) \varphi^*_{(\lambda)} (G^{(\lambda)}_{\boldsymbol{n}_{(\lambda)} \boldsymbol{\psi}_{(\lambda)}} + \Pi^{ab}_{(\lambda)} \mathcal{D}^{(\lambda)}_a \psi^{(\lambda)}_b). \tag{52}$$

Denoting $\mathscr{T} = \varphi^{-1}_{(\lambda)}(\mathscr{T}_{(\lambda)}) \subset \mathring{\mathscr{M}}$ as the inverse image of a constant time two-surface and similarly $\mathscr{B} = \varphi^{-1}_{(\lambda)}(\mathscr{B}_{(\lambda)}) \subset \mathring{\mathscr{M}}$ as the inverse image of the worldtube boundary (quasilocal frame) in the background manifold, and using the fact that the tensor transport commutes with contractions, the above can simply be written in the notation we have established as

$$\int_{\mathscr{T}_{\mathrm{f}} - \mathscr{T}_{\mathrm{i}}} (\varphi^*_{(\lambda)} \boldsymbol{\epsilon}_{\mathscr{T}_{(\lambda)}}) \Pi_{\tilde{\boldsymbol{u}} \boldsymbol{\psi}} = \int_{\Delta \mathscr{B}} (\varphi^*_{(\lambda)} \boldsymbol{\epsilon}_{\mathscr{B}_{(\lambda)}}) (G_{\boldsymbol{n} \boldsymbol{\psi}} + \Pi^{ab} \mathcal{D}_a \psi_b). \tag{53}$$

So far, we have been completely general. Now, let us restrict our attention to the momentum conservation law ($\boldsymbol{\psi} = -\boldsymbol{\phi} \in \mathscr{H}$) given by Eq. (53), and let us assume that we do not have any matter on $\Delta \mathscr{B}$ (hence, by the Einstein equation, $G_{\boldsymbol{n}\boldsymbol{\phi}}|_{\Delta \mathscr{B}} = \kappa T_{\boldsymbol{n}\boldsymbol{\phi}}|_{\Delta \mathscr{B}} = 0$), or even simply that any matter if present there is subdominant to the linear perturbation, i.e., $\boldsymbol{T}|_{\Delta \mathscr{B}} = \mathcal{O}(\lambda^2)$. The lhs then expresses

the change in momentum of the system (inside the worldtube interval in the perturbed spacetime) between some initial and final time slices; for notational ease, we will simply denote this by $\Delta \mathrm{p}^{(\boldsymbol{\phi})}$. [Note that we prefer to use typewriter font for the total quasilocal momentum, so as to avoid any confusion with matter four-momentum defined in the typical way from T_{ab} and traditionally labeled by P^a, as, e.g., in Eq. (8).] Then, inserting also the definition of the Brown-York tensor [Eq. (30)] on the rhs and replacing $\mathcal{D}$ with $\boldsymbol{\nabla}$ since it does not affect the contractions, Eq. (53) becomes

$$\Delta \mathrm{p}^{(\boldsymbol{\phi})} = \int_{\Delta\mathscr{B}} (\varphi^*_{(\lambda)} \boldsymbol{\epsilon}_{\mathscr{B}_{(\lambda)}}) \tau^{ab} \nabla_a \phi_b. \tag{54}$$

We claim, and will now demonstrate, that the $\mathcal{O}(\lambda)$ part of this always contains the GSF.

Let us consider Eq. (54) term by term. First, we have the transport—in this case, the pullback—under $\varphi_{(\lambda)}$ of the volume form of $\mathscr{B}_{(\lambda)}$. Now, we know that the pullback under a diffeomorphism of the volume form of a manifold is, in general, not simply the volume form of the inverse image of that manifold under the diffeomorphism. However, it is always true (see, e.g., Chapter 7 of Ref. [76]) that they are proportional, with the proportionality given by a smooth function called the Jacobian determinant and usually denoted by J. That is, in our case, we have $\varphi^*_{(\lambda)} \boldsymbol{\epsilon}_{\mathscr{B}_{(\lambda)}} = J \boldsymbol{\epsilon}_{\mathscr{B}}$, with $J \in C^\infty(\mathscr{B})$. In particular, this function is given by $J(p) = \det(T_p \varphi_{(\lambda)})$, $\forall\, p \in \mathscr{B}$, where $T_p \varphi_{(\lambda)} = (\varphi_{(\lambda)})_* : T_p \mathscr{B} \to T_{\varphi_{(\lambda)}} \mathscr{B}_{(\lambda)}$ is the pushforward, and the determinant is computed with respect to the volume forms $\boldsymbol{\epsilon}_{\mathscr{B}}(p)$ on $T_p \mathscr{B}$ and $\boldsymbol{\epsilon}_{\mathscr{B}_{(\lambda)}}(\varphi_{(\lambda)}(p))$ on $T_{\varphi_{(\lambda)}(p)} \mathscr{B}_{(\lambda)}$. Now, it is clear that we have $J = 1 + \mathcal{O}(\lambda)$, as $\varphi_{(0)}$ is simply the identity map. Therefore, we have

$$\varphi^*_{(\lambda)} \boldsymbol{\epsilon}_{\mathscr{B}_{(\lambda)}} = (1 + \mathcal{O}(\lambda)) \boldsymbol{\epsilon}_{\mathscr{B}}. \tag{55}$$

As for the other terms in the integrand of Eq. (54), we simply have

$$\tau^{ab} = \mathring{\tau}^{ab} + \lambda \delta \tau^{ab} + \mathcal{O}(\lambda^2), \tag{56}$$

$$\nabla_a \phi_b = \mathring{\nabla}_a \phi_b + \lambda \delta(\nabla_a \phi_b) + \mathcal{O}(\lambda^2). \tag{57}$$

Hence, we can see that there will be three contributions to the $\mathcal{O}(\lambda)$ rhs of Eq. (54). Respectively, from Eqs. (55)–(57), these are the $\mathcal{O}(\lambda)$ parts of the volume form pullback, which may not be easy to compute in practice; the Brown-York tensor $\boldsymbol{\tau}$, which may be computed from its definition [Eq. (30)]; and the derivative of the conformal Killing vector $\boldsymbol{\phi}$, which may be readily carried out and, as we will presently show, always contains the GSF. Thus, we denote this contribution to the $\mathcal{O}(\lambda)$ part of $\Delta \mathrm{p}^{(\boldsymbol{\phi})}$ as $\Delta \mathrm{p}^{(\boldsymbol{\phi})}_{\mathrm{self}}$,

$$\Delta \mathrm{p}^{(\boldsymbol{\phi})}_{\mathrm{self}} = \lambda \int_{\Delta\mathscr{B}} \boldsymbol{\epsilon}_{\mathscr{B}} \mathring{\tau}^{ab} \delta(\nabla_a \phi_b). \tag{58}$$

Now, we proceed with the computation of Eq. (58). In particular, let us consider the series expansion of Eq. (58) in the areal radius r of $\mathscr{B}$. This can be defined for any time slice by $r = (\frac{1}{4\pi} \int_{\mathscr{S}} \boldsymbol{\epsilon}_{\mathscr{S}})^{1/2}$, such that a constant r slice of $\mathring{\mathscr{M}}$ defines $\mathscr{B}$ (and $\boldsymbol{n} = M \mathring{\boldsymbol{\nabla}} r$ for some positive function M on $\mathscr{B}$). It has been shown [57] that the Brown-York tensor has, in general, the following expansion in r,

$$\mathring{\tau}^{ab} = \mathring{u}^a \mathring{u}^b \mathcal{E}_{\mathrm{vac}} - \mathring{\sigma}^{ab} \mathrm{P}_{\mathrm{vac}} + \mathcal{O}(r), \tag{59}$$

where

$$\mathcal{E}_{\mathrm{vac}} = -\frac{2}{\kappa r}, \tag{60}$$

$$\mathrm{P}_{\mathrm{vac}} = -\frac{1}{\kappa r} \tag{61}$$

are called the *vacuum energy* and *vacuum pressure*, respectively. Some remarks regarding these are warranted before we move on. In particular, these are terms which have sometimes been argued to play the role of "subtraction terms" (to be removed from the quasilocal energy-momentum tensor); see, e.g., Ref. [71]. From this point of view, the definition of the Brown-York tensor [Eq. (30)] may be regarded as carrying a certain amount of freedom, inasmuch as any freedom may be assumed to exist to define a "reference" action S_0 to be subtracted from the total (gravitational plus matter) action $S_{\mathrm{G+M}}$ in the variational principle discussed in Sec. II C. Such a subtraction of a reference action, while common practice in gravitational physics, has the sole function of shifting the numerical value of the action such that, ultimately, the numerical value of the Hamiltonian constructed from the modified action $S_{\mathrm{G+M}} - S_0$ may be interpreted as the Arnowitt-Deser-Misner (ADM) energy. However, this essentially amounts to a presumption that we are free to pick the zero of the energy—in other words, that the vacuum energy may be freely subtracted away without affecting the physics. Though we refrain from entering into much further detail here, it has been shown [51] that these vacuum terms, Eqs. (60) and (61), are in fact crucial for our conservation laws to yield physically reasonable answers and to make mathematical sense—evidencing that the vacuum energy/pressure should be taken seriously as having physically real significance. We will now lend further credibility to this by showing that they are precisely the energy (and pressure) associated with the momentum flux that are typically interpreted as the GSF. Actually, we argue in this paper that the term implicating the vacuum energy yields the standard form of the GSF, and the vacuum pressure term is novel in our analysis.

Now that we have an expansion [Eq. (59)] of $\mathring{\boldsymbol{\tau}}$ in r, let us consider the $\delta(\boldsymbol{\nabla}\boldsymbol{\phi})$ term. We see that

$$\delta(\nabla_a \phi_b) = \delta(\mathring{\nabla}_a \phi_b - C^d{}_{ab}\phi_d) = -\delta C^c{}_{ab}\phi_c. \quad (62)$$

Collecting all of our results so far—inserting Eqs. (59)–(62) into Eq. (58)—we thus get

$$\Delta \mathrm{p}^{(\boldsymbol{\phi})}_{\mathrm{self}} = \lambda \frac{2}{\kappa} \int_{\Delta\mathscr{B}} \boldsymbol{\epsilon}_{\mathscr{B}} \frac{1}{r} (\mathring{u}^a \mathring{u}^b - 2\mathring{\sigma}^{ab}) \delta C^c{}_{ab} \phi_c + \mathcal{O}(r). \quad (63)$$

Let us now look at the contractions in the integrand. For the first (energy) term, inserting the connection coefficient (50), we have by direct computation

$$\mathring{u}^a \mathring{u}^b \delta C^c{}_{ab}\phi_c = \mathring{g}^{cd}\left(\mathring{\nabla}_a h_{bd} - \frac{1}{2}\mathring{\nabla}_d h_{ab}\right)\mathring{u}^a \mathring{u}^b \phi_c \quad (64)$$

$$= -F^c[\boldsymbol{h}; \mathring{\boldsymbol{u}}]\phi_c, \quad (65)$$

where the functional $\boldsymbol{F}$ is precisely the GSF four-vector functional defined in the introduction [Eq. (5)], and to write the final equality, we have used the orthogonality property $\phi_{\mathring{\boldsymbol{u}}} = 0$. Thus, we see that this is indeed the term that yields the GSF. For the second (pressure) term in Eq. (63), we similarly obtain by direct computation

$$\mathring{\sigma}^{ab}\delta C^c{}_{ab}\phi_c = 2\wp^c[\boldsymbol{h}; \mathring{\boldsymbol{\sigma}}]\phi_c, \quad (66)$$

where, in expressing the rhs, it is convenient to define a general functional of two (0,2)-tensors similar to the GSF functional:

$$\wp^c[\boldsymbol{H}; \boldsymbol{S}] = \frac{1}{2}\mathring{g}^{cd}\left(\mathring{\nabla}_a H_{bd} - \frac{1}{2}\mathring{\nabla}_d H_{ab}\right) S^{ab}. \quad (67)$$

We call this novel term the gravitational self-pressure force.

Now, we can collect all of the above and insert them into (63). Before writing down the result, it is convenient to define a total functional $\boldsymbol{\mathcal{F}}$ as the sum of $\boldsymbol{F}$ and $\wp$,

$$\mathscr{F}^a[\boldsymbol{h}; \mathring{\boldsymbol{u}}] = F^a[\boldsymbol{h}; \mathring{\boldsymbol{u}}] + \wp^a[\boldsymbol{h}; \mathring{\boldsymbol{\sigma}}]. \quad (68)$$

We refer to this as the *extended GSF functional*. Note that for $\boldsymbol{\mathcal{F}}$ we write only the functional dependence on $\boldsymbol{h}$ and $\mathring{\boldsymbol{u}}$ since the two-metric $\mathring{\boldsymbol{\sigma}}$ is determined uniquely by $\mathring{\boldsymbol{u}}$. With this, and setting the perturbation parameter to unity, Eq. (63) becomes

$$\boxed{\Delta \mathrm{p}^{(\boldsymbol{\phi})}_{\mathrm{self}} = -\frac{1}{4\pi}\int_{\Delta\mathscr{B}} \boldsymbol{\epsilon}_{\mathscr{B}} \frac{1}{r} \boldsymbol{\phi} \cdot \boldsymbol{\mathcal{F}}[\boldsymbol{h}; \mathring{\boldsymbol{u}}] + \mathcal{O}(r).} \quad (69)$$

This is to be compared with Gralla's formula [40] discussed in the Introduction, Eq. (7). While the equivalence thereto is immediately suggestive based on the general form of our result, we have to do a bit more work to show that indeed Eq. (69), both on the lhs and the rhs, recovers—though in general will, evidently at least from our novel gravitational self-pressure force, also have extra terms added to—Eq. (7). We leave this task to the following section, the purpose of which is to consider in detail the application of our conservation law formulation to a concrete example of a perturbative family of spacetimes defined for a self-force analysis, namely, the Gralla-Wald family.

Concordantly, we emphasize that the result above [Eq. (69)] holds for any family of perturbed manifolds $\{\mathscr{M}_{(\lambda)}\}_{\lambda\geq 0}$ and is completely independent of the internal description of our system, i.e., the worldtube interior. In other words, what we have just demonstrated—provided only that one accepts a quasilocal notion of energy-momentum—is that the (generalized) GSF is a completely generic perturbative effect in GR for localized systems; it arises as a linear-order contribution of any spacetime perturbation to the momentum flux of a system in the limit where its areal radius is small.

This view of the self-force may cast fresh conceptual light on the old and seemingly arcane problem of deciphering its physical origin and meaning. In particular, recall the common view that the GSF is caused by the back-reaction of the "mass" of a small object upon its own motion. Yet, what we have seen here is that it is actually the vacuum mass, or vacuum energy that is responsible for the GSF. We may still regard the effect as a "backreaction," in the sense that it is the boundary metric perturbations of the system—the $\boldsymbol{h}$ on $\mathscr{B}$—which determine its momentum flux, but the point is that this flux is inexorably present and given by Eq. (69) regardless of where exactly this $\boldsymbol{h}$ is coming from. Presumably, the dominant part of $\boldsymbol{h}$ would arise from the system itself—if we further assume that the system itself is indeed what is being treated perturbatively by the family $\{\mathscr{M}_{(\lambda)}\}_{\lambda\geq 0}$, as is the case with typical self-force analyses—but in principle, $\boldsymbol{h}$ can comprise absolutely any perturbations; i.e., its physical origin does not even have to be from inside the system.

In this way, we may regard the GSF as a completely geometrical, purely general-relativistic backreaction of the mass (and pressure) of the spacetime vacuum—not of the object inside—upon the motion of a localized system (i.e., its momentum flux). This point of view frees us from having to invoke such potentially ambiguous notions as "mass ratios" (in a two-body system for example), let alone "Coulombian m/r fields," to make basic sense of self-force effects. They simply—and always—happen from the interaction of the vacuum with any boundary perturbation, and are dominant if that boundary is not too far out.

IV. APPLICATION TO THE GRALLA-WALD APPROACH TO THE GRAVITATIONAL SELF-FORCE

In this section, we will consider in detail the application of our ideas to a particular approach to the self-force, that is

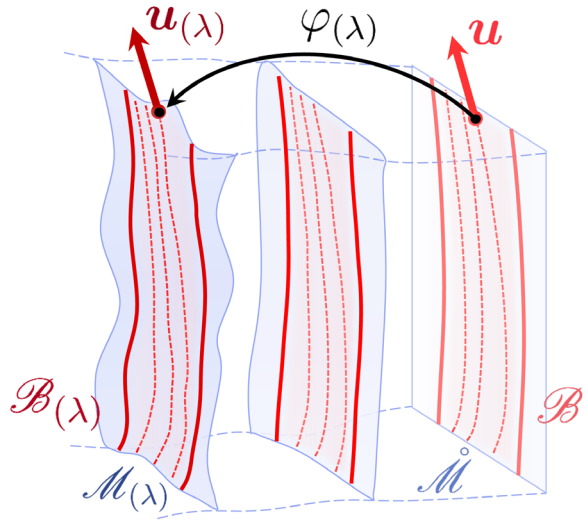

FIG. 7. Representation of a one-parameter family of QFs $\{(\mathscr{B}_{(\lambda)};\boldsymbol{u}_{(\lambda)})\}_{\lambda\geq 0}$ embedded correspondingly in a family of spacetimes $\{\mathscr{M}_{(\lambda)}\}_{\lambda\geq 0}$.

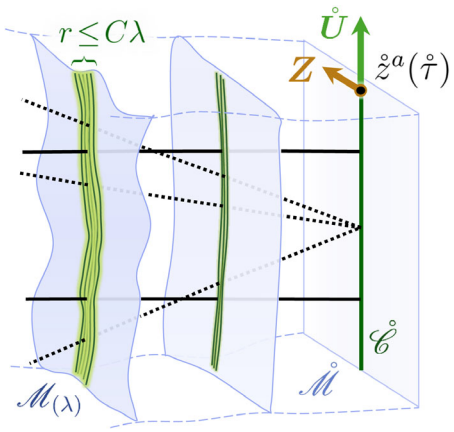

FIG. 8. Representation of the Gralla-Wald family of spacetimes $\{\mathscr{M}_{(\lambda)}\}_{\lambda\geq 0}$. (This is an adaptation of Fig. 1 of Ref. [33].) The lined green region that "fills in" $\mathscr{M}_{(\lambda)}$ for $r\leq C\lambda$ is the small object which scales down to zero size and mass in the background $\mathring{\mathscr{M}}$. The solid black lines represent taking the ordinary limit (the far away view where the motion appears reduced to a worldline), and the dashed black lines represent the scaled limit (the close by view where the rest of the Universe appears pushed away to infinity). The worldline $\mathring{\mathscr{C}}$, which can be proven to be a geodesic, is parametrized by $\mathring{z}^a(\mathring{\tau})$ and has four-velocity $\mathring{\boldsymbol{U}}$. The deviation vector $\boldsymbol{Z}$ on $\mathring{\mathscr{C}}$ is used for formulating the first-order correction to the motion.

to say, a particular specification of $\{(\mathscr{M}_{(\lambda)},\boldsymbol{g}_{(\lambda)})\}$ via a few additional assumptions aimed at encoding the notion of a small object being "scaled down" to zero "size" and mass as $\lambda\to 0$. In other words, we now identify the perturbation (which has up to this point been treated completely abstractly) defined by $\{(\mathscr{M}_{(\lambda)},\boldsymbol{g}_{(\lambda)})\}$ as actually being that caused by the presence of the small object; that could mean regular matter (in particular, a compact object such as a neutron star) or a black hole.

The assumptions (on $\{\boldsymbol{g}_{(\lambda)}\}$) that we choose to work with here are those of the approach of Gralla and Wald [33]. Certainly, the application of our perturbed quasilocal conservation laws could just as well be carried out in the context of any other self-force analysis—such as, e.g., the self-consistent approximation of Pound [77] (the mathematical correspondence of which to the Gralla-Wald approach has, in any case, been shown in Ref. [78]).

Our motivation for starting with the Gralla-Wald approach in particular is twofold. On the one hand, it furnishes a mathematically rigorous and physically clear picture (which we show in Fig. 8)—arguably more so than any other available GSF treatment—of what it means to scale down a small object to zero size and mass (or, equivalently, of perturbing any spacetime by the presence of an object with small size and mass—we will be more precise momentarily). On the other hand, it is within this approach that the formula for the GSF has been obtained (in Ref. [40]) as a closed two-surface (small two-sphere) integral around the object (in lieu of evaluating the GSF at a spacetime point identified as the location of the object), in the form of the Gralla angle averaging formula [Eq. (7)]—with which our extended GSF formula (69) is to be compared.

In Sec. IV A, we provide an overview of the assumptions and consequences of the Gralla-Wald approach to the GSF. Afterward, in Sec. IV B, we describe the general embedding of rigid quasilocal frames in the Gralla-Wald family of spacetimes, and then in Sec. IV C, we describe their detailed construction in the background spacetime in this family. Having established this, we then proceed to derive equations of motion in two ways. In particular, we carry out the analysis with two separate choices of rigid quasilocal frames ("frames of reference"): first, inertially with the point-particle approximation of the moving object in the background in Sec. IV D and, second, inertially with the object itself in the perturbed spacetime in Sec. IV E.

A. Gralla-Wald approach to the GSF

The basic idea of Gralla and Wald [33] for defining a family $\{(\mathscr{M}_{(\lambda)},\boldsymbol{g}_{(\lambda)})\}_{\lambda\geq 0}$ such that $\lambda>0$ represents the inclusion of perturbations generated by a small object is the following one. One begins by imposing certain smoothness conditions on $\{\boldsymbol{g}_{(\lambda)}\}_{\lambda\geq 0}$ corresponding to the existence of certain limits of each $\boldsymbol{g}_{(\lambda)}$. In particular, two limits are sought corresponding intuitively to two limiting views of the system: first, a view from "far away" from which the "motion" of the (extended but localized) object reduces to a worldline and, second, a view from "close by" the object from which the rest of the Universe (and, in particular, the MBH it might be orbiting as in an EMRI) looks "pushed away" to infinity. A third requirement must be added to

this, namely, that both of these limiting pictures nonetheless coexist in the same spacetime; i.e., the two limits are smoothly related (or, in other words, there is no pathological behavior when taking these limits along different directions). While in principle this may sound rather technical, one can actually motivate each of these conditions with very sensible physical arguments as we shall momentarily elaborate further upon. From them, Gralla and Wald have shown [33] that it is possible to derive a number of consequences, including geodesic motion in the background at zeroth order and the MiSaTaQuWa equation [34,35] for the GSF at first order in λ.

Let us now be more precise. Let $\{(\mathscr{M}_{(\lambda)}, \boldsymbol{g}_{(\lambda)})\}_{\lambda\geq 0}$ be a perturbative one-parameter family of spacetimes as in the previous section. We assume that $\{\boldsymbol{g}_{(\lambda)}\}_{\lambda\geq 0}$ satisfies the following conditions, depicted visually in Fig. 8:

(i) *Existence of an "ordinary limit".*—There exist coordinates $\{x^\alpha\}$ in $\mathscr{M}_{(\lambda)}$ such that $g^{(\lambda)}_{\beta\gamma}(x^\alpha)$ is jointly smooth in (λ, x^α) for $r > C\lambda$ where $C > 0$ is a constant and $r = (x_i x^i)^{1/2}$. For all $\lambda \geq 0$ and $r > C\lambda$, $\boldsymbol{g}_{(\lambda)}$ is a vacuum solution of the Einstein equation. Furthermore, $\overset{\circ}{g}_{\beta\gamma}(x^\alpha)$ is smooth in x^α including at $r = 0$, and the curve $\overset{\circ}{\mathscr{C}} = \{r = 0\} \subset \overset{\circ}{\mathscr{M}}$ is timelike.

(ii) *Existence of a "scaled limit".*—For all t_0, define the "scaled coordinates" $\{\bar{x}^\alpha\} = \{\bar{t}, \bar{x}^i\}$ by $\bar{t} = (t - t_0)/\lambda$ and $\bar{x}^i = x^i/\lambda$. Then, the "scaled metric" $\bar{g}^{(\lambda)}_{\bar{\beta}\bar{\gamma}}(t_0; \bar{x}^\alpha) = \lambda^{-2} g^{(\lambda)}_{\bar{\beta}\bar{\gamma}}(t_0; \bar{x}^\alpha)$ is jointly smooth in $(\lambda, t_0; \bar{x}^\alpha)$ for $\bar{r} = r/\lambda > C$.

(iii) *Uniformity condition.*—Define $A = r$, $B = \lambda/r$, and $n^i = x^i/r$. Then, each $g^{(\lambda)}_{\beta\gamma}(x^\alpha)$ is jointly smooth in (A, B, n^i, t).

Mathematically, the first two conditions respectively ensure the existence of an appropriate Taylor expansion (in r and λ) of the metric in a "far zone" (on length scales comparable with the mass of the MBH in an EMRI, $r \sim M$) and a "near zone" (on length scales comparable with the mass of the object, $r \sim m$). Meanwhile, the third is simply a consistency requirement ensuring the existence of a "buffer zone" ($m \ll r \ll M$) where both expansions are valid. (This idea is in many ways similar to the method of "matched asymptotic expansions" [34].)

From a physical point of view, what is happening in the first ("ordinary") limit is that the body is shrinking down to a worldline $\overset{\circ}{\mathscr{C}}$ with its mass (understood as defining the perturbation) going to zero at least as fast as its radius. (As we increase the perturbative parameter λ from zero, the radius is not allowed to grow faster than linearly with λ; viewed conversely, this condition ensures that the object does not collapse to a black hole if it was not one already before reaching the point-particle limit.) In the second (scaled) limit, the object is shrinking down to zero size in an asymptotically self-similar manner (its mass is proportional to its size, and its "shape" is not changing). Finally, the uniformity condition ensures that there are no "bumps of curvature" in the one-parameter family. (Essentially, this guarantees that there are no inconsistencies in evaluating the limits along different directions.)

From these assumptions alone, Gralla and Wald [33] are able to derive the following consequences:

(a) *Background motion.*—The worldline $\overset{\circ}{\mathscr{C}}$ is a geodesic in $\overset{\circ}{\mathscr{M}}$; writing its parametrization in terms of proper time $\overset{\circ}{\tau}$ as $\overset{\circ}{\mathscr{C}} = \{\overset{\circ}{z}{}^a(\overset{\circ}{\tau})\}_{\overset{\circ}{\tau}\in\mathbb{R}}$ and denoting its four-velocity by $\overset{\circ}{U}{}^a = \mathrm{d}\overset{\circ}{z}{}^a(\overset{\circ}{\tau})/\mathrm{d}\overset{\circ}{\tau}$, this means that

$$\overset{\circ}{\nabla}_{\overset{\circ}{\boldsymbol{U}}} \overset{\circ}{\boldsymbol{U}} = 0. \tag{70}$$

(b) *Background scaled metric.*—$\overset{\circ}{\bar{\boldsymbol{g}}}$ is stationary and asymptotically flat.

(c) *First-order field equation.*—At $\mathcal{O}(\lambda)$, the Einstein equation is sourced by the matter energy-momentum tensor of a point particle $\boldsymbol{T}^{\mathrm{PP}}$ supported on $\overset{\circ}{\mathscr{C}}$; i.e., the field equation is

$$\delta G_{ab}[\boldsymbol{h}] = \kappa T^{\mathrm{PP}}_{ab}, \tag{71}$$

where

$$T^{\mathrm{PP}}_{ab} = m \int_{\overset{\circ}{\mathscr{C}}} \mathrm{d}\overset{\circ}{\tau}\, \overset{\circ}{U}_a(\overset{\circ}{\tau}) \overset{\circ}{U}_b(\overset{\circ}{\tau}) \delta_4(x^c - z^c(\overset{\circ}{\tau})). \tag{72}$$

Here, m is a constant along $\overset{\circ}{\mathscr{C}}$ and is interpreted as representing the mass of the object—or, more precisely, the mass of the point particle which approximates the object in the background. (This is a subtle point that should be kept in mind, and which will be better elucidated in our analysis further on.)

(d) *First-order equation of motion.*—At $\mathcal{O}(\lambda)$, the correction to the motion in the Lorenz gauge—corresponding to the choice of a certain gauge vector $\mathsf{L} \in T\mathcal{N}$ defined by the condition

$$\overset{\circ}{\nabla}{}^b \left(h^{\mathsf{L}}_{ab} - \frac{1}{2} h^{\mathsf{L}} \overset{\circ}{g}_{ab} \right) = 0, \tag{73}$$

where $h = \mathrm{tr}(\boldsymbol{h})$—is given by the MiSaTaQuWa equation [34,35],

$$\overset{\circ}{\nabla}_{\overset{\circ}{\boldsymbol{U}}} \overset{\circ}{\nabla}_{\overset{\circ}{\boldsymbol{U}}} Z^a = -\overset{\circ}{E}_b{}^a Z^b + F^a[\boldsymbol{h}^{\mathrm{tail}}; \overset{\circ}{\boldsymbol{U}}], \tag{74}$$

where $\overset{\circ}{E}_b{}^a = \overset{\circ}{R}_{cbd}{}^a \overset{\circ}{U}{}^c \overset{\circ}{U}{}^d$ is the electric part of the Weyl tensor and $\boldsymbol{h}^{\mathrm{tail}}$ is a tail integral of the retarded Green's functions of $\boldsymbol{h}$. The above is an equation for a four-vector $\boldsymbol{Z}$ called the deviation vector; the lhs is the acceleration associated therewith, and the rhs is a geodesic deviation term plus the GSF. This deviation vector is defined on $\overset{\circ}{\mathscr{C}}$ and represents the first-order

correction needed to move off $\mathring{\mathscr{C}}$ and onto the worldline representing the center of mass of the perturbed spacetime, defined as in the Hamiltonian analysis of Regge and Teitelboim [41].

Let us make a few comments on these results, specifically concerning *(a)* and *(c)*. On the one hand, it is quite remarkable that geodesic motion can be recovered as a consequence[10] of this analysis-i.e., without having to posit it as an assumption—just from smoothness properties (existence of appropriate limits) of our family of metrics $\{\boldsymbol{g}_{(\lambda)}\}$, and on the other, this analysis offers sensible meaning to the usual "delta function cartoon" (ubiquitous in essentially all self-force analyses) of the matter stress-energy-momentum tensor describing the object in the background spacetime. The point is that the description of the object is completely arbitrary inside the region that is not covered by the smoothness conditions of the family $\{\boldsymbol{g}_{(\lambda)}\}$, i.e., for $r \leq C\lambda$ when $\lambda > 0$. (Indeed, this region can be filled in even with exotic matter, e.g., failing to satisfy the dominant energy condition, or a naked singularity, as long as a well-posed initial value formulation exists.) Regardless of what this description is, the smoothness conditions essentially ensure that its "reduction" to $\mathring{\mathscr{M}}$ (or, more precisely, the transport of any effect thereof with respect to the family $\{\boldsymbol{g}_{(\lambda)}\}$) simply becomes that of a point particle sourcing the field equation at $\mathcal{O}(\lambda)$. In this way, the background "point-particle cartoon" is justified as the simplest possible idealization of a small object.

What we are going to do, essentially, is to accept consequences *(a)–(c)* [in fact, we will not even explicitly need *(b)*], the proofs of which do not rely upon any further limiting conditions such as a restriction of the perturbative gauge, and to obtain, using our perturbed momentum conservation law, a more general version of the EoM, i.e., consequence *(d)*. For the latter, Gralla and Wald [33] instead rely on the typical but laborious Hadamard expansion techniques of DeWitt and Brehme [44], wherein the "mass dipole moment" of the object is set to zero. It is possible [41] to have such a notion in a well-defined Hamiltonian sense by virtue of *(b)*. While mathematically rigorous and conducive to obtaining the correct known form of the MiSaTaQuWa equation, their derivation and final result suffer not only from the limitation of having to fix the perturbative gauge but also from the (as we shall see, potentially avoidable) technical complexity of arriving at the final answer—including the evaluation of $\mathbf{h}^{\mathrm{tail}}$ (or otherwise taking recourse to a regularization procedure).

The link between this approach and our conservation law derivation of the EoM which we are about to carry out is established by the work of Gralla [40], who discovered that Eq. (74) can be equivalently written as

$$\mathring{\nabla}_{\mathring{\boldsymbol{U}}}\mathring{\nabla}_{\mathring{\boldsymbol{U}}}Z^a = -\mathring{E}_b{}^a Z^b + \frac{1}{4\pi}\lim_{r\to 0}\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}F^a[\mathring{\boldsymbol{U}},\boldsymbol{h}]. \tag{75}$$

Here, the GSF term $\boldsymbol{F}[\boldsymbol{h}^{\mathrm{tail}};\mathring{\boldsymbol{U}}]$ in the MiSaTaQuWa equation [Eq. (74)] is substituted by an integral expression—an average over the angles—of $\boldsymbol{F}$. In particular (as, strictly speaking, one cannot define integrals of vectors as such), this is evaluated by using the exponential map based on $\mathring{\mathscr{C}}$ to associate a flat metric, in terms of which the integration is performed over a two-sphere of radius r, $\mathbb{S}^2_r$, with $\boldsymbol{\epsilon}_{\mathbb{S}^2}$ denoting the volume form of $\mathbb{S}^2$.

Observe that, here, the functional dependence of $\boldsymbol{F}$ is on $\boldsymbol{h}$ itself (and not on $\boldsymbol{h}^{\mathrm{tail}}$ or any sort of regularized $\boldsymbol{h}$) and for this reason is referred to as the "bare" GSF. Moreover, this formula is actually valid in a wider class of gauges than just the Lorenz gauge, in particular, it holds in what are referred to as parity-regular gauges [40]. We refrain from entering here into the technical details of exactly how such gauges are defined, except to say that the eponymous "parity condition" that they need to satisfy has its ultimate origin in the Hamiltonian analysis of Regge and Teitleboim [41] and is imposed so as to make certain Hamiltonian definitions—and in particular for Gralla's analysis [40], the Hamiltonian center of mass—well defined. These, however, are not limitations of our quasilocal formalism, where we know how to define energy-momentum notions more generally than any Hamiltonian approach. Thus, in our result, there will be *no restriction* on the perturbative gauge. This may constitute a great advantage, as the parity-regular gauge class—though an improvement from being limited to the Lorenz gauge in formulating the EoM—still excludes entire classes of perturbative gauges convenient for formulating black hole perturbation theory (e.g., the Regge-Wheeler gauge in Schwarzschild-Droste) and hence for carrying out practical EMRI calculations.

We proceed to apply our quasilocal analysis to the Gralla-Wald family of spacetimes, beginning with a general setup in this family of rigid quasilocal frames.

B. General setup of rigid quasilocal frames in the Gralla-Wald family

Let $(\mathscr{B}_{(\lambda)};\boldsymbol{u}_{(\lambda)})$ be a quasilocal frame in $(\mathscr{M}_{(\lambda)},\boldsymbol{g}_{(\lambda)},\boldsymbol{\nabla}_{(\lambda)})$, for any $\lambda > 0$, constructed just as described in Sec. II: with unit four-velocity $\boldsymbol{u}_{(\lambda)}$, unit normal $\boldsymbol{n}_{(\lambda)}$, induced metric $\boldsymbol{\gamma}_{(\lambda)}$, and so on. Using the fact that the tensor transport is linear and commutes with tensor products, we can compute the transport (in the five-dimensional stacked manifold $\mathscr{N} = \mathscr{M}_{(\lambda)} \times \mathbb{R}^{\geq}$ used in our perturbative setup, as in Sec. III A) of any geometrical quantity of interest to the background. For example,

$$\gamma_{ab} = \varphi^*_{(\lambda)}\gamma^{(\lambda)}_{ab} \tag{76}$$

$$= \varphi^*_{(\lambda)}(g^{(\lambda)}_{ab} - n^{(\lambda)}_a n^{(\lambda)}_b) \tag{77}$$

[10]See again footnote 2 and the references mentioned therein for more on this topic.

$$= g_{ab} - n_a n_b \tag{78}$$

$$= \overset{\circ}{\gamma}_{ab} + \lambda\delta\gamma_{ab} + \mathcal{O}(\lambda^2), \tag{79}$$

where

$$\overset{\circ}{\gamma}_{ab} = \overset{\circ}{g}_{ab} - \overset{\circ}{n}_a \overset{\circ}{n}_b, \tag{80}$$

$$\delta\gamma_{ab} = h_{ab} - 2\overset{\circ}{n}_{(a}\delta n_{b)}. \tag{81}$$

Similarly,

$$\sigma_{ab} = \overset{\circ}{\sigma}_{ab} + \lambda\delta\sigma_{ab} + \mathcal{O}(\lambda^2), \tag{82}$$

where

$$\overset{\circ}{\sigma}_{ab} = \overset{\circ}{\gamma}_{ab} + \overset{\circ}{u}_a \overset{\circ}{u}_b, \tag{83}$$

$$\delta\sigma_{ab} = \delta\gamma_{ab} + 2\overset{\circ}{u}_{(a}\delta u_{b)}. \tag{84}$$

Now, let us assume that $(\mathscr{B}_{(\lambda)}; \boldsymbol{u}_{(\lambda)})$ is a rigid quasilocal frame, meaning that the congruence defining it has a vanishing symmetrized strain rate tensor in $\mathscr{M}_{(\lambda)}$,

$$\theta^{(\lambda)}_{(ab)} = 0. \tag{85}$$

Let $\mathscr{B} = \varphi^{-1}_{(\lambda)}(\mathscr{B}_{(\lambda)})$ be the inverse image of $\mathscr{B}_{(\lambda)}$ in the background $\overset{\circ}{\mathscr{M}}$, with $\boldsymbol{u} = \varphi^{*}_{(\lambda)}\boldsymbol{u}_{(\lambda)} = \overset{\circ}{\boldsymbol{u}} + \lambda\delta\boldsymbol{u} + \mathcal{O}(\lambda^2)$ giving the transport of the quasilocal observers' four-velocity, $\boldsymbol{n} = \overset{\circ}{\boldsymbol{n}} + \lambda\delta\boldsymbol{n} + \mathcal{O}(\lambda^2)$ the unit normal, and so on. In other words, $(\mathscr{B}; \boldsymbol{u})$ is the background mapping of the perturbed congruence $(\mathscr{B}_{(\lambda)}; \boldsymbol{u}_{(\lambda)})$, and so will itself constitute a congruence (in the background), i.e., a quasilocal frame defined by a two-parameter family of worldlines with unit four-velocity $\boldsymbol{u}$ in $\overset{\circ}{\mathscr{M}}$.

However, although $(\mathscr{B}_{(\lambda)}; \boldsymbol{u}_{(\lambda)})$ is a rigid quasilocal frame in $\mathscr{M}_{(\lambda)}$, $(\mathscr{B}; \boldsymbol{u})$ is not in general a rigid quasilocal frame in $\overset{\circ}{\mathscr{M}}$ (with respect to the background metric $\overset{\circ}{\boldsymbol{g}}$). One can see this easily as follows. Let $\boldsymbol{\vartheta} \in \mathcal{J}^0{}_2(\overset{\circ}{\mathscr{M}})$ be the strain rate tensor of $(\mathscr{B}; \boldsymbol{u})$, so that it is given by

$$\vartheta_{ab} = \sigma_{ca}\sigma_{bd}\overset{\circ}{\nabla}{}^c u^d. \tag{86}$$

The rhs is a series in λ, owing to the fact that $\boldsymbol{u}$ (and therefore $\boldsymbol{\sigma}$, the two-metric on the space $\mathscr{H}$ orthogonal to $\boldsymbol{u}$ in $\mathscr{B}$) is transported from a perturbed congruence in $\mathscr{M}_{(\lambda)}$. Upon expansion, we obtain

$$\vartheta_{ab} = \overset{\circ}{\vartheta}_{ab} + \lambda\delta\vartheta_{ab} + \mathcal{O}(\lambda^2), \tag{87}$$

where

$$\overset{\circ}{\vartheta}_{ab} = \overset{\circ}{\sigma}_{c(a}\overset{\circ}{\sigma}_{b)d}\overset{\circ}{\nabla}{}^c \overset{\circ}{u}{}^d \tag{88}$$

is just the strain rate tensor of the background congruence—i.e., the congruence defined by $\overset{\circ}{\boldsymbol{u}}$—and

$$\delta\vartheta_{ab} = 2\overset{\circ}{\sigma}{}^{(c}{}_{(a}\delta\sigma^{d)}{}_{b)}\overset{\circ}{\nabla}_c \overset{\circ}{u}_d + \overset{\circ}{\sigma}{}^c{}_{(a}\overset{\circ}{\sigma}_{b)d}\overset{\circ}{\nabla}_c \delta u^d \tag{89}$$

is the first-order term in λ. Note that we are abusing our established notation slightly in writing Eq. (87), as there exists no $\boldsymbol{\vartheta}_{(\lambda)}$ in $\mathscr{M}_{(\lambda)}$ of which the transport (to $\overset{\circ}{\mathscr{M}}$) yields such a series expansion; instead, $\boldsymbol{\vartheta}$ is defined directly on $\overset{\circ}{\mathscr{M}}$ (relative to the metric $\overset{\circ}{\boldsymbol{g}}$) as the strain rate tensor of a congruence with four-velocity $\boldsymbol{u}$—which itself contains the expansion in λ.

Now, let us compute the transport of the rigidity condition on $(\mathscr{B}_{(\lambda)}; \boldsymbol{u}_{(\lambda)})$ [Eq. (85)] to $\overset{\circ}{\mathscr{M}}$: we have

$$0 = \varphi^{*}_{(\lambda)}\theta^{(\lambda)}_{(ab)} \tag{90}$$

$$= \varphi^{*}_{(\lambda)}(\sigma^{(\lambda)}_{c(a}\sigma^{(\lambda)}_{b)d}\nabla^c_{(\lambda)}u^d_{(\lambda)}) \tag{91}$$

$$= \sigma_{c(a}\sigma_{b)d}\nabla^c u^d \tag{92}$$

$$= \overset{\circ}{\theta}_{(ab)} + \lambda\delta\theta_{(ab)} + \mathcal{O}(\lambda^2), \tag{93}$$

where

$$\overset{\circ}{\theta}_{(ab)} = \overset{\circ}{\vartheta}_{(ab)}, \tag{94}$$

$$\delta\theta_{(ab)} = \delta\vartheta_{(ab)} + \overset{\circ}{\sigma}{}^c{}_{(a}\overset{\circ}{\sigma}_{b)d}\delta C^d{}_{ce}\overset{\circ}{u}{}^e. \tag{95}$$

Since $0 = \theta^{(\lambda)}_{(ab)}$ identically in $\mathscr{M}_{(\lambda)}$ [as we demand that $(\mathscr{B}_{(\lambda)}; \boldsymbol{u}_{(\lambda)})$ is a rigid quasilocal frame], Eq. (93) must vanish order by order in λ. That implies, in particular, that the zeroth-order congruence (defined by $\overset{\circ}{\boldsymbol{u}}$) is a rigid quasilocal frame, and that the symmetrized strain rate tensor of the background-mapped perturbed congruence (defined by $\boldsymbol{u}$) is given by

$$\vartheta_{(ab)} = -\lambda\overset{\circ}{\sigma}{}^c{}_{(a}\overset{\circ}{\sigma}_{b)d}\delta C^d{}_{ce}\overset{\circ}{u}{}^e + \mathcal{O}(\lambda^2). \tag{96}$$

This tells us that the deviation from rigidity of $(\mathscr{B}; \boldsymbol{u})$ in $\overset{\circ}{\mathscr{M}}$ occurs only at $\mathcal{O}(\lambda)$ (and, in particular, is caused by the same perturbed connection coefficient term that is responsible for the GSF). In other words, we can treat $(\mathscr{B}; \boldsymbol{u})$ as a rigid quasilocal frame at zeroth order. This zeroth-order congruence actually makes up a different worldtube

boundary $\mathring{\mathscr{B}} \neq \mathscr{B}$ in $\mathring{\mathscr{M}}$, i.e., one defined by a congruence with four-velocity $\mathring{\boldsymbol{u}} \neq \boldsymbol{u}$ in general. Clearly, for a rigid quasilocal frame with a small areal radius r constructed around a worldline $\mathscr{G}$ in $\mathring{\mathscr{M}}$ with four-velocity $\boldsymbol{U}_{\mathscr{G}}$, we would simply have $\mathring{\boldsymbol{u}} = \boldsymbol{U}_{\mathscr{G}}$ (where the rhs is understood to be transported off $\mathscr{G}$ and onto $\mathring{\mathscr{B}}$ via the exponential map), and $\mathring{\boldsymbol{\sigma}} = r^2\mathfrak{S}$; i.e., it is the metric of $\mathbb{S}^2_r$. This is the most trivial possible rigid quasilocal frame: at any instant of time, a two-sphere worth of quasilocal observers moving with the same four-velocity as the point at its center (parametrizing the given worldline).

At first order, the equation $0 = \delta\theta_{(ab)}$ can be regarded as the constraint on the linear perturbations ($\delta\boldsymbol{u}$) in the motion of the quasilocal observers in terms of the metric perturbations guaranteeing that the perturbed congruence is rigid in the perturbed spacetime. (So, presumably, going to nth order in λ would yield equations for every term up to the nth-order piece of the motion of the quasilocal observers, $\delta^n\boldsymbol{u}$.)

Now, recall the momentum conservation law for rigid quasilocal frames, Eq. (39). This holds for $(\mathscr{B}_{(\lambda)}; \boldsymbol{u}_{(\lambda)})$ in $\mathscr{M}_{(\lambda)}$. Just as we did in the previous section with the general conservation law, we can use $\varphi_{(\lambda)}$ to turn this into an equation in $\mathring{\mathscr{M}}$:

$$\Delta \mathrm{p}^{(\boldsymbol{\phi})} = -\int_{\Delta\mathscr{B}} \varphi^*_{(\lambda)} \boldsymbol{\epsilon}_{\mathscr{B}_{(\lambda)}} (\mathcal{E}\alpha_{\boldsymbol{\phi}} + 2\nu\epsilon_{ab}\mathcal{P}^a\boldsymbol{\phi}^b + \mathrm{P}\boldsymbol{D}\cdot\boldsymbol{\phi}). \tag{97}$$

Let us now further assume that we can ignore the Jacobian determinant (discussed in the previous section) as well as the shift $\boldsymbol{v}$ of the quasilocal observers (relative to constant time surfaces). Then, dividing the above equation by Δt, where t represents the adapted time coordinate on $\mathscr{B}$, and taking the $\Delta t \to 0$ limit, we get the time rate of change of the momentum,

$$\dot{\mathrm{p}}^{(\boldsymbol{\phi})} = -\int_{\mathscr{I}} \boldsymbol{\epsilon}_{\mathscr{I}} N\tilde{\gamma}(\mathcal{E}\alpha_{\boldsymbol{\phi}} + 2\nu\epsilon_{ab}\mathcal{P}^a\boldsymbol{\phi}^b + \mathrm{P}\boldsymbol{D}\cdot\boldsymbol{\phi}), \tag{98}$$

where $\dot{\mathrm{p}}^{(\boldsymbol{\phi})} = \mathrm{d}\mathrm{p}^{(\boldsymbol{\phi})}/\mathrm{d}t$, and we must keep in mind that the derivative is with respect to the adapted time on (the inverse image on the background of) our congruence.

C. Detailed construction of background rigid quasilocal frames

Let $\mathscr{G}$ be any timelike worldline in $\mathring{\mathscr{M}}$. Any background metric $\mathring{\boldsymbol{g}}$ on $\mathring{\mathscr{M}}$ in a neighborhood of $\mathscr{G}$ admits an expression in Fermi normal coordinates [36,50], which we label by $\{X^\alpha\} = \{T = X^0, X^I\}_{I=1}^3$, as a power series in the areal radius. Denoting by $A_K(T)$ and $W_K(T)$ the proper acceleration and proper rate of rotation of the spatial axes (triad) along $\mathscr{G}$ (as functions of the proper time T along $\mathscr{G}$), respectively, this is given by

$$\mathring{g}_{00} = -(1 + A_K X^K)^2 + R^2 W_K W_L P^{KL} - \mathring{R}_{0K0L}X^K X^L + \mathcal{O}(R^3), \tag{99}$$

$$\mathring{g}_{0J} = \epsilon_{JKL}W^K X^L - \frac{2}{3}\mathring{R}_{0KJL}X^K X^L + \mathcal{O}(R^3), \tag{100}$$

$$\mathring{g}_{IJ} = \delta_{IJ} - \frac{1}{3}\mathring{R}_{IKJL}X^K X^L + \mathcal{O}(R^3), \tag{101}$$

where $R^2 = \delta_{IJ}X^I X^J$ is the square of the radius in these coordinates (not the square of the Ricci scalar) and $P^{KL} = \delta^{KL} - X^K X^L/R^2$ projects vectors perpendicular to the radial direction X^I/R. Here, we have to remember that the Riemann tensor $\mathring{R}_{IJKL}$ (along with $\boldsymbol{A}$ and $\boldsymbol{W}$) is understood to be evaluated on $\mathscr{G}$.

For all cases that we will be interested in, we will ignore the possibility of rotation, so we set $W_I = 0$ from now on.

Let us now assume that our background rigid quasilocal frame $(\mathring{\mathscr{B}}; \mathring{\boldsymbol{u}})$ is constructed around $\mathscr{G}$; that is to say, into this coordinate system there is embedded a two-parameter family of worldlines representing a topological two-sphere worth of observers, i.e., a fibrated timelike worldtube $\mathring{\mathscr{B}}$ surrounding $\mathscr{G}$. This may be conveniently described, as detailed in Sec. II B, by defining a new set of coordinates $\{x^\alpha\} = \{t, r, x^{\mathrm{i}}\}_{\mathrm{i}=1}^2$ given simply by the adapted coordinates $\{t, x^{\mathrm{i}}\}_{\mathrm{i}=1}^2$ on $\mathring{\mathscr{B}}$ supplemented with a radial coordinate r. Then, denoting $\{x^{\mathrm{i}}\} = \{\theta, \phi\}$, we introduce, as done in previous calculations with rigid quasilocal frames in Fermi normal coordinates [57], the following coordinate transformation,

$$T(t, r, \theta, \phi) = t + \mathcal{O}(r^2, \mathcal{R}), \tag{102}$$

$$X^I(t, r, \theta, \phi) = r r^I(\theta, \phi) + \mathcal{O}(r^2, \mathcal{R}), \tag{103}$$

where

$$r^I(\theta, \phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \tag{104}$$

are the standard direction cosines of a radial unit vector in spherical coordinates in $\mathbb{R}^3$ and $\mathcal{R}$ here represents the order of the perturbations of the quasilocal frame away from the round two-sphere due to the background curvature effects. In particular, for rigid quasilocal frames, we know that this is in fact simply the order of the Riemann tensor on $\mathscr{G}$, i.e., $\mathring{R}_{IJKL} = \mathcal{O}(\mathcal{R})$. Thus, one may ultimately desire to take $\mathcal{O}(\mathcal{R})$ effects into account for a full calculation, but for the moment—since, in principle, this $\mathcal{R}$ is unrelated to λ and

we can assume it to be subdominant thereto—we simply omit them. Thus, we can simply take $\mathring{\mathscr{I}} = \mathbb{S}^2_r$, and we can assume that there is no shift, so that $\tilde{\gamma} = 1$.

Applying the coordinate transformation in Eqs. (102) and (103) to the background metric given by Eqs. (99)–(101) with $\boldsymbol{W} = 0$, and then using all of the definitions that we have established so far, it is possible to obtain by direct computation all of the quantities appearing in the integrand of the conservation law [Eq. (98)] as series in r. We display the results only up to leading order in r, including the possibility of setting $\boldsymbol{A} = 0$:

$$\mathring{N} = 1 + rA_I r^I + \frac{1}{2} r^2 \mathring{E}_{IJ} r^I r^J + \mathcal{O}(r^3), \tag{105}$$

$$\begin{aligned} \mathring{\mathcal{E}} &= \mathcal{E}_{\text{vac}} + \mathcal{O}(r) \\ &= -\frac{2}{\kappa r} + \mathcal{O}(r), \end{aligned} \tag{106}$$

$$\mathring{\alpha}_{\mathfrak{i}} = rA_I \mathfrak{B}^I_{\mathfrak{i}} + r^2 (\mathring{E}_{IJ} - A_I A_J) \mathfrak{B}^I_{\mathfrak{i}} r^J + \mathcal{O}(r^3), \tag{107}$$

$$\mathring{\nu} = -r \mathring{B}_{IJ} r^I r^J + \mathcal{O}(r^2), \tag{108}$$

$$\mathring{\mathcal{P}}_{\mathfrak{i}} = -\frac{1}{\kappa} r^2 \mathring{B}_{IJ} \mathfrak{R}^I_{\mathfrak{i}} r^J + \mathcal{O}(r^3), \tag{109}$$

$$\begin{aligned} \mathring{\mathrm{P}} &= \mathrm{P}_{\text{vac}} - \frac{1}{\kappa} A_I r^I + \mathcal{O}(r) \\ &= -\frac{1}{\kappa r} - \frac{1}{\kappa} A_I r^I + \mathcal{O}(r). \end{aligned} \tag{110}$$

Here, $\mathring{E}_{IJ} = \mathring{C}_{0I0J}|_{\mathscr{G}}$ and $\mathring{B}_{IJ} = \frac{1}{2} \epsilon_I{}^{KL} \mathring{C}_{0JKL}|_{\mathscr{G}}$ are, respectively, the electric and magnetic parts of the Weyl tensor evaluated on $\mathscr{G}$. Also, $\mathfrak{B}^I_{\mathfrak{i}} = \partial_{\mathfrak{i}} r^I$ and $\mathfrak{R}^I_{\mathfrak{i}} = \epsilon^{\mathbb{S}^2}_{\mathfrak{i}}{}^{\mathfrak{j}} \mathfrak{B}^I_{\mathfrak{j}}$ are, respectively, the boost and rotation generators of $\mathbb{S}^2$. See Appendix A for more technical details on this. We remind the reader that $\mathcal{E}_{\text{vac}}$ and P_{vac} are, respectively, the vacuum energy and pressure, Eqs. (60) and (61), respectively.

The way to proceed is now clear: we expand Eq. (98) as a series in λ,

$$\dot{\mathfrak{p}}^{(\boldsymbol{\phi})} = (\dot{\mathfrak{p}}^{(\boldsymbol{\phi})})_{(0)} + \lambda \delta \dot{\mathfrak{p}}^{(\boldsymbol{\phi})} + \mathcal{O}(\lambda^2), \tag{111}$$

using the zeroth-order parts of the various terms written above. We need only to specify the worldline $\mathscr{G}$ in $\mathring{\mathscr{M}}$ about which we are carrying out the Fermi normal coordinate expansion (in r). We will consider two cases: $\mathscr{G} = \mathring{\mathscr{C}}$ (the geodesic, such that $\mathscr{B}$ is inertial with the point particle in $\mathring{\mathscr{M}}$) and $\mathscr{G} = \mathscr{C}$ (an accelerated worldline such that $\mathscr{B}_{(\lambda)}$ is inertial with the object in $\mathscr{M}_{(\lambda)}$, i.e., it is defined by a constant $r > C\lambda$ in $\mathscr{M}_{(\lambda)}$). These will give us equivalent descriptions of the dynamics of the system, from two different "points of view," or (quasilocal) frames of reference.

Before entering into the calculations, we can simplify things further by remarking that the zeroth-order expansions in Eqs. (105)–(110) will always make the twist (ν) term in the conservation law [Eq. (98)] appear at $\mathcal{O}(r)$ or higher, both in $(\dot{\mathfrak{p}}^{(\boldsymbol{\phi})})_{(0)}$ and $\delta \dot{\mathfrak{p}}^{(\boldsymbol{\phi})}$, regardless of our choice of $\mathscr{G}$. Hence, we can safely ignore it, as we are interested (at least for this work) only in the part of the conservation law which is zeroth order in r. Thus, we simply work with

$$\dot{\mathfrak{p}}^{(\boldsymbol{\phi})} = -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2 N (\mathcal{E} \alpha_{\boldsymbol{\phi}} + \mathrm{P} \boldsymbol{D} \cdot \boldsymbol{\phi}). \tag{112}$$

Into this, we furthermore have to insert the multipole expansion of the conformal Killing vector $\boldsymbol{\phi}$ given by Eq. (41). We correspondingly write

$$\dot{\mathfrak{p}}^{(\boldsymbol{\phi})} = \sum_{\ell \in \mathbb{N}} \dot{\mathfrak{p}}^{(\boldsymbol{\phi}_\ell)}, \tag{113}$$

such that for any $\ell \in \mathbb{N}$, we have

$$\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_\ell)} = -\Phi^{I_1 \cdots I_\ell} \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r N (\mathcal{E} \alpha_{\mathfrak{i}} + \mathrm{P} D_{\mathfrak{i}}) D^{\mathfrak{i}} \left(\prod_{n=1}^{\ell} r_{I_n} \right). \tag{114}$$

Explicitly, the first two terms are

$$\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})} = -\Phi^I \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r N (\mathcal{E} \alpha_{\mathfrak{i}} + \mathrm{P} D_{\mathfrak{i}}) \mathfrak{B}^{\mathfrak{i}}_I, \tag{115}$$

$$\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=2})} = -2\Phi^{IJ} \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r N (\mathcal{E} \alpha_{\mathfrak{i}} + \mathrm{P} D_{\mathfrak{i}}) (\mathfrak{B}^{\mathfrak{i}}_I r_J). \tag{116}$$

D. Equation of motion inertial with the background point particle

Let $\mathscr{G} = \mathring{\mathscr{C}}$. Then, $\boldsymbol{A} = 0$. We will take this to be the case for the rest of this subsection—corresponding, as discussed, to a rigid quasilocal frame the inverse image in the background of which is inertial with the point-particle approximation of the moving object in the background spacetime. This situation is displayed visually in Fig. 9.

Let us first compute the zeroth-order (in λ) part of $\dot{\mathfrak{p}}^{(\boldsymbol{\phi})}$. Inserting (105)–(110) into the zeroth-order part of (115) and (116), and making use of the various properties in Appendix A, we find by direct computation:

$$(\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})})_{(0)} = \mathcal{O}(r^2), \tag{117}$$

$$(\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=2})})_{(0)} = \mathcal{O}(r^2). \tag{118}$$

We provide the steps of the calculation in Appendix B.

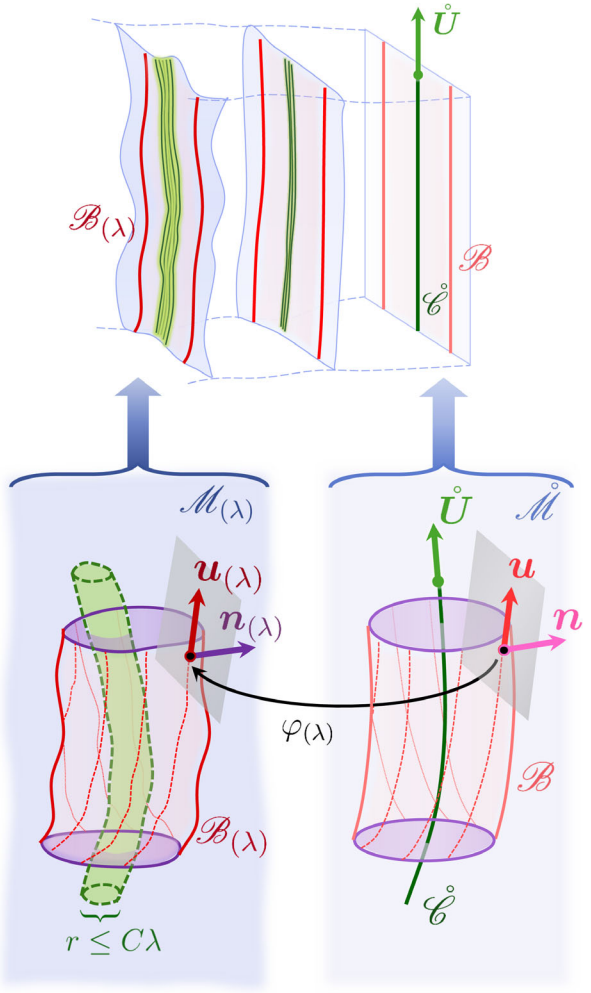

FIG. 9. A family of rigid quasilocal frames $\{(\mathscr{B}_{(\lambda)}; \boldsymbol{u}_{(\lambda)})\}$ embedded in the Gralla-Wald family of spacetimes $\{\mathscr{M}_{(\lambda)}\}$ such that the inverse image of any such perturbed quasilocal frame in the background is inertial with the point-particle approximation of the moving object, i.e., is centered on the geodesic $\mathring{\mathscr{C}}$.

Let us now compute the $\mathcal{O}(\lambda)$, $\ell = 1$ part of $\dot{\mathfrak{p}}^{(\boldsymbol{\phi})}$, i.e., the $\mathcal{O}(\lambda)$ part of Eq. (115), which as usual we denote by $\delta\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})}$. One can see that this will involve contributions from five $\mathcal{O}(\lambda)$ terms, respectively containing δN, $\delta\mathcal{E}$, $\delta\boldsymbol{\alpha}$, δP, and $\delta\boldsymbol{D}$. For convenience, we will use the notation $(\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_\ell)}_{(Q)})_{(n)}$ to indicate the term of $\delta^n(\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_\ell)})$ that is linear in Q, for any ℓ, n. Thus, we write

$$\delta\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})} = \sum_{Q\in\{\delta N,\delta\mathcal{E},\delta\boldsymbol{\alpha},\delta\mathrm{P},\delta\boldsymbol{D}\}} \delta\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(Q)}. \tag{119}$$

All of the computational steps are again in Appendix B. We find

$$\delta\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta N)} = -\frac{2}{\kappa}\Phi_I \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2}\delta N r^I + \mathcal{O}(r^2). \tag{120}$$

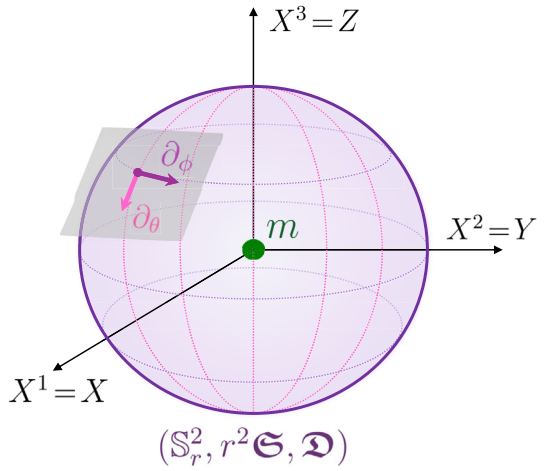

FIG. 10. An instantaneous rigid quasilocal frame $(\mathbb{S}^2_r, r^2\mathfrak{S}, \mathfrak{D})$ (where $\mathfrak{S}$ and $\mathfrak{D}$, respectively, are the metric and derivative compatible with the unit two-sphere) inertial with the background point particle. This means that the latter is located at the center of our Fermi normal coordinate system.

If δN does not vary significantly over $\mathbb{S}^2_r$, the $\mathcal{O}(r^0)$ part of the above would be negligible owing to the fact that $\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^I = 0$.

Next, let us consider the $\delta\mathcal{E}$ and δP terms. For this, we find it useful to depict the instantaneous quasilocal frame $(\mathbb{S}^2_r, r^2\boldsymbol{S}, \boldsymbol{D})$ embedded in a constant-time three-slice of $\mathring{\mathscr{M}}$ in Fig. 10.

The $\delta\mathcal{E}$ term can be easily determined by realizing that in our current choice of quasilocal frame, the only background matter is the point particle which is always at the center of our present coordinate system; i.e., it is always on $\mathring{\mathscr{C}}$ (on which we are here centering our Fermi normal coordinates). Interpreting the constant m as in the Gralla-Wald approach [33] to be the mass of this point particle, this simply means that

$$\delta\mathcal{E} = \frac{m}{4\pi r^2}, \tag{121}$$

so that when this is integrated (as a surface energy density) over $\mathbb{S}^2_r$, we simply recover the mass: $\int_{\mathbb{S}^2_r} r^2\boldsymbol{\epsilon}_{\mathbb{S}^2}\delta\mathcal{E} = m$. We remark that, by definition, it is possible to express the quasilocal energy as $\mathcal{E} = u^a u^b \tau_{ab} = -\frac{1}{\kappa}k$ with $k = \boldsymbol{\sigma}:\boldsymbol{\Theta}$ the trace of the two-dimensional boundary extrinsic curvature. Notice that the integral of this over a closed two-surface in the $r \to \infty$ limit is in fact the same as the usual ADM definition of the mass/energy; thus, $\delta\mathcal{E} = -\frac{1}{\kappa}\delta k$, and so it makes sense to interpret m as the ADM mass of the object. So, now, using Eq. (121), we can find that the $\delta\mathcal{E}$ contribution to $\delta\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})}$ is also at most quadratic in r:

$$\delta\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\mathcal{E})} = \mathcal{O}(r^2). \tag{122}$$

To compute the δP term, we now employ the useful identity in Eq. (40), which tells us that

$$\delta \mathrm{P} = \frac{1}{2}\delta\mathcal{E} - \frac{1}{\kappa}\delta a_{\boldsymbol{n}}. \tag{123}$$

Using this, into which we insert the $\delta\mathcal{E}$ from Eq. (121), we find that the δP contribution to $\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}$ is at most quadratic in r as well,

$$\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta \mathrm{P})} = \mathcal{O}(r^2). \tag{124}$$

Note that the above results may in fact be higher order in r than quadratic. We have only explicitly checked that they vanish up to linear order inclusive.

Finally, we are left with the $\delta\boldsymbol{\alpha}$ and $\delta\boldsymbol{D}$ contributions to $\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}$. By direct computation, it is possible to show that their sum is in fact precisely what we have referred to as the extended GSF in our general analysis of the preceding section, i.e., it is the $\ell = 1$ part of Eq. (69),

$$\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\boldsymbol{\alpha})} + \delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\boldsymbol{D})} = \frac{\mathrm{d}}{\mathrm{d}t}(\Delta \mathrm{p}^{(\boldsymbol{\phi}_{\ell=1})}_{\mathrm{self}}). \tag{125}$$

In particular, they respectively contribute the usual GSF (from $\delta\boldsymbol{\alpha}$) and the gravitational self-pressure force (from $\delta\boldsymbol{D}$).

Thus, we have found that the total $\mathcal{O}(\lambda)$, $\ell = 1$ part of the momentum time rate of change is given at leading (zeroth) order in r by nothing more than the generalized GSF. In other words,

$$\boxed{\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})} = -\Phi_I \mathrm{F}^I + \mathcal{O}(r)}, \tag{126}$$

where we have defined

$$\mathrm{F}^I = -\frac{2}{\kappa}\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2}\mathfrak{S}^{\mathrm{i}\mathrm{j}}\mathfrak{B}^I_{\mathrm{i}}\mathcal{F}_{\mathrm{j}}[\boldsymbol{h};\mathring{\boldsymbol{u}}] + \mathcal{O}(r). \tag{127}$$

Without loss of generality, let us now pick $\Phi^I = (0,0,1)$ to be the unit vector in the Cartesian $X^3 = Z$ direction, and denote its corresponding conformal Killing vector as $\boldsymbol{\phi}_{\ell=1} = \boldsymbol{\phi}^Z_{\ell=1}$. (Alternately, pick the Z axis to be oriented along Φ^I.) We know $\mathfrak{S}^{\mathrm{i}\mathrm{j}}\mathfrak{B}^Z_{\mathrm{j}} = (-1/\sin\theta, 0)$; moreover, by the coordinate transformation $\mathcal{F}_{\mathrm{i}} = (\partial x^J/\partial x^{\mathrm{i}})\mathcal{F}_J$, we have $\mathcal{F}_\theta = \cos\theta(\cos\phi\mathcal{F}_X + \sin\phi\mathcal{F}_Y) - \sin\theta\mathcal{F}_Z$. Inserting these into Eq. (126), we get

$$\begin{aligned}\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}^Z_{\ell=1})} = &-\frac{2}{\kappa}\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}\mathcal{F}_Z[\boldsymbol{h};\mathring{\boldsymbol{u}}] \\ &+\frac{2}{\kappa}\int_{\mathbb{S}^2_r}\mathrm{d}\theta\wedge\mathrm{d}\phi\cos\theta(\cos\phi\mathcal{F}_X + \sin\phi\mathcal{F}_Y).\end{aligned} \tag{128}$$

The first line is precisely in the form of the GSF term from the Gralla formula, Eq. (7) [40], except here in the integrand we have (the Z-component of) our extended GSF $\boldsymbol{\mathcal{F}}$ [Eq. (11)]: the usual GSF $\boldsymbol{F}$ (the only self-force term in Gralla's formula) plus our self-pressure term, $\wp$. The second line contains additional terms involving the extended GSF in the other two (Cartesian) spatial directions. Notice, however, that $\int_{\mathbb{S}^2_r}\mathrm{d}\theta\wedge\mathrm{d}\phi\cos\theta\cos\phi = 0 = \int_{\mathbb{S}^2_r}\mathrm{d}\theta\wedge\mathrm{d}\phi\cos\theta\sin\phi$, so if $\mathcal{F}_X$ and $\mathcal{F}_Y$ do not vary significantly over $\mathbb{S}^2_r$, their contribution will be subdominant to that of $\mathcal{F}_Z$.

Thus, we have shown that our EoM (126) always contains Gralla's angle average of the bare (usual) GSF. However, the form of (126) (expressing the perturbative change in the quasilocal momentum) still cannot be *directly* compared, as such, with Gralla's EoM (75) (expressing the change in a deviation vector representing the perturbative correction to the motion). In the following subsection, we clarify the correspondence by repeating the calculation using a quasilocal frame inertial with the moving extended object in the perturbed spacetime (rather than with the geodesic in the background, as here). Furthermore, we conjecture that a careful imposition of the parity condition on the perturbative gauge—of which we have made no explicit use so far—would make the contribution from our self-pressure term vanish, but a detailed proof is required and remains to be carried out.

E. Equation of motion inertial with the moving object in the perturbed spacetime

Now, let $\mathscr{G} = \mathscr{C} \neq \mathring{\mathscr{C}}$ (so $\boldsymbol{A} \neq 0$ in general) such that the quasilocal frame $(\mathscr{B};\boldsymbol{u})$ centered on $\mathscr{C}$ (in $\mathring{\mathscr{M}}$) is the inverse image of the rigid quasilocal frame $(\mathscr{B}_{(\lambda)};\boldsymbol{u}_{(\lambda)})$ defined by $r = C\lambda + \varepsilon = \text{const}$, $\forall\,\varepsilon > 0$, in $\mathscr{M}_{(\lambda)}$. The meaning of the r coordinate in the latter is as given in the Gralla-Wald assumptions (Sec. IV A). This situation is displayed in Fig. 11.

We now proceed to calculate, in the same way as we did for the point-particle-inertial case, the various terms in the expansion of the momentum conservation law, Eqs. (115) and (116). At zeroth order, we obtain

$$(\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})})_{(0)} = \mathcal{O}(r^2), \tag{129}$$

$$(\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=2})})_{(0)} = \mathcal{O}(r^2). \tag{130}$$

The steps of all these computations are again shown in Appendix B.

Let us now compute the $\mathcal{O}(\lambda)$, $\ell = 1$ part of $\dot{\mathrm{p}}^{(\boldsymbol{\phi})}$. First, we find that $\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta N)}$ is the same as in the point-particle-inertial case, so if δN does not vary significantly over $\mathbb{S}^2_r$, the $\mathcal{O}(r^0)$ part thereof is negligible.

Next, let us look at the $\delta\mathcal{E}$ and δP parts. Again, it is useful to consider in this case the visual depiction of the instantaneous quasilocal frame, shown in Fig. 12.

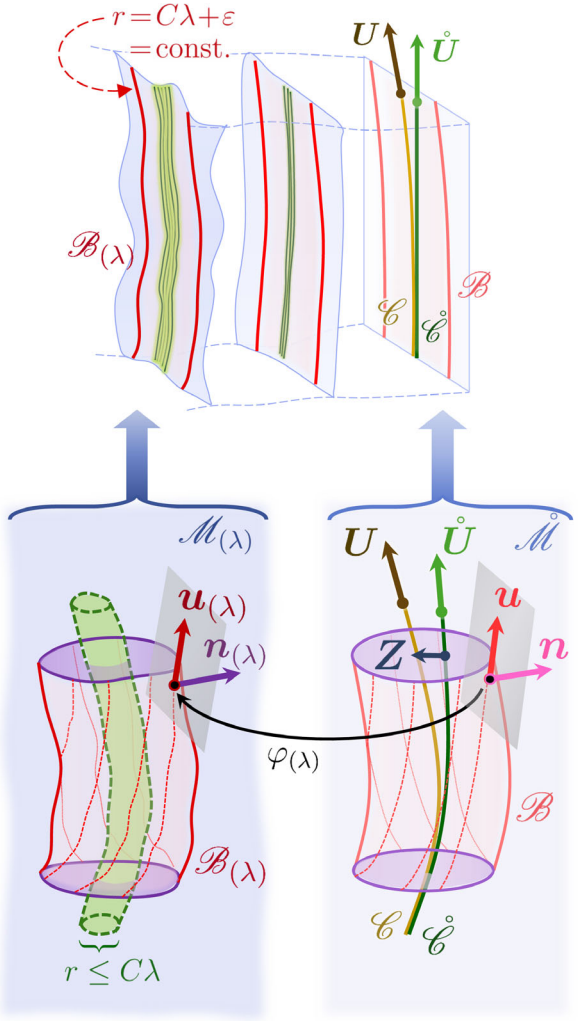

FIG. 11. A family of rigid quasilocal frames $\{(\mathscr{B}_{(\lambda)}; \boldsymbol{u}_{(\lambda)})\}$ embedded in the Gralla-Wald family of spacetimes $\{\mathscr{M}_{(\lambda)}\}$ inertial with the moving object in $\mathscr{M}_{(\lambda)}$. This means that $\mathscr{B}_{(\lambda)}$ is defined by the constancy of the Gralla-Wald r coordinate in $\mathscr{M}_{(\lambda)}$, for any $r > C\lambda$. Thus, the inverse image $\mathscr{B}$ of $\mathscr{B}_{(\lambda)}$ in the background $\mathring{\mathscr{M}}$ is centered, in general, not on the geodesic $\mathring{\mathscr{C}}$ followed by the point-particle background approximation of the object but on some timelike worldline $\mathscr{C} \neq \mathring{\mathscr{C}}$, with four-velocity $\boldsymbol{U} \neq \mathring{\boldsymbol{U}}$, which may be regarded as an approximation on $\mathring{\mathscr{M}}$ of the true motion of the object in $\mathscr{M}_{(\lambda)}$. Between $\mathring{\mathscr{C}}$ and $\mathscr{C}$, there is a deviation vector $\boldsymbol{Z}$, which is to be compared with the deviation vector (correction to the motion) in the Gralla-Wald approach.

In this case, the particle (delta function) will not be at the center of our coordinate system but instead displaced in some direction ρ^I relative thereto. Nonetheless, we know that this displacement must itself be $\mathcal{O}(\lambda)$ which means that it will only contribute $\mathcal{O}(\lambda)$ corrections to the $\delta\mathcal{E}$ having m exactly at the center; i.e., we have

$$\delta\mathcal{E} = \frac{m}{4\pi(X^I - \lambda\rho^I)(X_I - \lambda\rho_I)} = \frac{m}{4\pi r^2} + \mathcal{O}(\lambda), \tag{131}$$

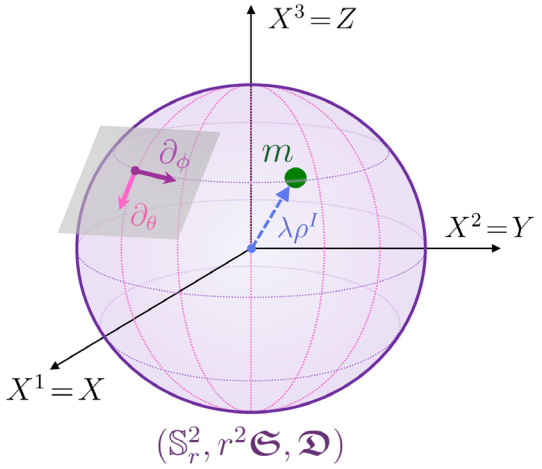

FIG. 12. An instantaneous rigid quasilocal frame $(\mathbb{S}^2_r, r^2\mathfrak{S}, \mathfrak{D})$ (where $\mathfrak{S}$ and $\mathfrak{D}$, respectively, are the metric and derivative compatible with the unit two-sphere) inertial with the moving object in the perturbed spacetime. This means that the point-particle approximation of this object in the background spacetime is not located at the center of our Fermi normal coordinate system. Instead, it is displaced in some direction ρ^I, which must be $\mathcal{O}(\lambda)$.

and as before, $\delta\mathrm{P} = \frac{1}{2}\delta\mathcal{E} - \frac{1}{\kappa}\delta a_{\boldsymbol{n}}$. Using these, we find

$$\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\mathcal{E})} = -\frac{2}{3}m\Phi_I A^I + \mathcal{O}(r), \tag{132}$$

$$\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\mathrm{P})} = +\frac{1}{3}m\Phi_I A^I + \mathcal{O}(r), \tag{133}$$

with the steps shown in Appendix B. Thus,

$$\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\mathcal{E})} + \delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\mathrm{P})} = -\frac{1}{3}m\Phi_I A^I + \mathcal{O}(r). \tag{134}$$

Meanwhile, we still have, exactly as in the point-particle-inertial case,

$$\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\boldsymbol{\alpha})} + \delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\boldsymbol{D})} = -\Phi_I \mathrm{F}^I + \mathcal{O}(r). \tag{135}$$

Now, by construction, we know that here $\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})} = 0$, as we are inertial with the moving object (in the "actual" spacetime $\mathscr{M}_{(\lambda)}$). Thus, summing the above and equating them to zero, we get

$$0 = \Phi_I(-mA^I - 3\mathrm{F}^I) + \mathcal{O}(r). \tag{136}$$

Since Φ^I is arbitrary, we thus get the EoM

$$mA^I = -3\mathrm{F}^I \tag{137}$$

in the $r \to 0$ limit.

Finally, to cast this EoM into the same form as Gralla-Wald [33], i.e., in terms of a deviation vector $\boldsymbol{Z}$ on $\mathring{\mathscr{C}}$ rather

than in terms of the proper acceleration $\boldsymbol{A}$ of $\mathscr{C}$, we use the generalized deviation equation (as the name suggests, the deviation equation between arbitrary worldlines, not necessarily geodesics), Eq. (37) of Ref. [79]. In our case, this reads $\lambda\ddot{Z}^I = \lambda A^I - \lambda Z^J \overset{\circ}{E}{}^I{}_J + \mathcal{O}(\lambda^2)$. Combining this with Eq. (137), we finally recover the $\mathcal{O}(\lambda)$ EoM

$$\boxed{\lambda m\ddot{Z}^I = -3\lambda \mathrm{F}^I - \lambda \overset{\circ}{E}{}^I{}_J Z^J + \mathcal{O}(\lambda^2, r)}. \tag{138}$$

Note that the factor of 3 multiplying the self-force term is in fact present in the EoM in Gralla's Appendix B, that is Eq. (B3) of Ref. [40]. The latter, in this case, expresses the time evolution not of the deviation vector itself but of the change in this deviation vector due to a gauge transformation, possibly including extra terms in case that transformation is out of the parity-regular class. We conjecture that a detailed analysis of the precise correspondence between our deviation vector definition and that of Gralla-Wald (which, while encoding the same intuitive notion of a perturbative correction to the motion, may not be completely identical in general), together with a relation of their gauge transformation properties, would make it possible to relate these EoM exactly.

V. CONCLUSIONS

In this paper, we have used quasilocal conservation laws to develop a novel formulation of self-force effects in general relativity, one that is independent of the choice of the perturbative gauge and applicable to any perturbative scheme designed to describe the correction to the motion of a localized object. In particular, we have shown that the correction to the motion of any finite spatial region, due to any perturbation of any spacetime metric, is dominated when that region is small (i.e., at zeroth order in a series expansion in its areal radius) by an extended gravitational self-force; this is the standard gravitational self-force term known up to now plus a new term, not found in previous analyses and attributable to a gravitational pressure effect with no analog in Newtonian gravity, which we have dubbed the gravitational self-pressure force. Mathematically, we have found that the total change in momentum $\Delta \mathrm{p}^{(\boldsymbol{\phi})} = \mathrm{p}^{(\boldsymbol{\phi})}_{\mathrm{final}} - \mathrm{p}^{(\boldsymbol{\phi})}_{\mathrm{initial}}$ between an initial and final time of any (gravitational plus matter) system subject to any metric perturbation $\boldsymbol{h}$ is given, in a direction determined by a conformal Killing vector $\boldsymbol{\phi}$ (see Sec. II D), by the following flux through the portion of the quasilocal frame (worldtube boundary) $(\mathscr{B}; \overset{\circ}{\boldsymbol{u}})$ delimited thereby,

$$\Delta \mathrm{p}^{(\boldsymbol{\phi})} = -\frac{c^4}{4\pi G}\int_{\Delta\mathscr{B}} \boldsymbol{\epsilon}_{\mathscr{B}} \frac{1}{r}\boldsymbol{\phi}\cdot\boldsymbol{\mathcal{F}}[\boldsymbol{h}; \overset{\circ}{\boldsymbol{u}}] + \mathcal{O}(r), \tag{139}$$

where we have restored units, r is the areal radius, and $\boldsymbol{F}$ is the extended self-force functional. In particular, $\boldsymbol{\mathcal{F}} = \boldsymbol{F} + \wp$ where $\boldsymbol{F}$ is the usual bare self-force [determined by the functional in Eq. (5)], and $\wp$ is our novel self-pressure force [determined by the functional in Eq. (67)].

The most relevant practical application of the self-force is in the context of modeling EMRIs. Ideally, one would like to compute the correction to the motion at the location of the moving object (SCO). Yet, once a concrete perturbative procedure is established, the latter usually ends up being described by a distribution (Dirac delta function), rendering such a computation ill defined unless additional tactics (typically in the form of regularizations or Green's functions methods) are introduced. However, if one takes a step back from the exact point denoting the location of the "particle" (the distributional support), and instead considers a flux around it, any singularities introduced in such a model are avoided by construction.

We have, moreover, shown that our formulation, when applied in the context of one particular and very common approach to the self-force—namely, that of Gralla and Wald [33]—yields equations of motion of the same form as those known up to now; in particular, they always contain, in the appropriate limit, the angle average self-force term of Gralla [40]. We conjecture that a more rigorous study of these equations of motion and their gauge transformation properties would prove their exact correspondence under appropriate conditions.

We would like here to offer a concluding discussion on our results in this paper in Sec. V A, as well as an outlook toward future work in Sec. V B.

A. Discussion of results

From a physical point of view, our approach offers a fresh and conceptually clear perspective on the basic mechanism responsible for the emergence of self-force effects in general relativity. In particular, we have demonstrated that the self-force may be regarded as nothing more than the manifestation of a physical flux of gravitational momentum passing through the boundary enclosing the small moving object. This gravitational momentum, and gravitational stress-energy-momentum in general, cannot be defined locally in general relativity. As we have argued at length in this paper, such notions must instead be defined quasilocally, i.e., as a boundary rather than as a volume densities. This is why the self-force appears mathematically as a boundary integral around the moving object [Eq. (69)], dominant in the limit where the areal radius is small.

The interpretation of the physical meaning of the self-force as a consequence of conservation principles leads to many interesting implications. As we have seen, the mass of the moving object—e.g., the mass m of the SCO in the EMRI problem—seems to have nothing to do *fundamentally* with the general existence of a self-force effect. Indeed, according to our analysis, the self-force is in fact generically present as a correction to the motion—and dominant when the moving region is small—whenever one

has *any* perturbation $\boldsymbol{h}$ to the spacetime metric that is nonvanishing on the boundary of the system.

The usual way to understand the gravitational self-force up to now has been to regard it as a backreaction of m on the metric, i.e., on the gravitational field, and thus in turn upon its own motion through that field. Schematically, one thus imagines that the linear correction to the motion is "linear in m" (or more generally, that the full correction is an infinite series in m), i.e., that it has the form $\delta\dot{\mathfrak{p}} \sim m\delta\mathfrak{a}$, with a "perturbed acceleration" $\delta\mathfrak{a}$ determined by $\boldsymbol{h}$ (according to some perturbative prescription) causing a correction to the momentum $\delta\dot{\mathfrak{p}}$ by a (linear) coupling to the mass m.

Our analysis, instead, shows that this momentum correction $\delta\dot{\mathfrak{p}}$ actually arises fundamentally in the schematic form

$$\delta\dot{\mathfrak{p}} \sim \mathcal{E}_{\text{vac}}\delta\mathfrak{a} + \mathrm{P}_{\text{vac}}\delta D, \tag{140}$$

where $\mathcal{E}_{\text{vac}}$ and P_{vac} are the vacuum energy and pressure [Eqs. (60) and (61), respectively] and $\delta\mathfrak{a}$ and δD are perturbed acceleration and gradient terms determined by $\boldsymbol{h}$. Thus, it is the vacuum energy (or mass) and vacuum pressure, not the mass of the moving object, which are responsible for the backreaction that produces self-force corrections.

Certainly, the metric perturbation $\boldsymbol{h}$ on the system boundary determining the perturbed acceleration and gradient terms in (140) may in turn be sourced by a "small mass" present in the interior of the system. In fact, if indeed the system is small, there may well be little physical reason for expecting that (the dominant part of) $\boldsymbol{h}$ would come for anything *other than* the presence of the small system itself. Concordantly, the aim of any concrete self-force analysis is to prescribe exactly *how* $\boldsymbol{h}$ is sourced thereby. Nevertheless, the correction (139) is valid *regardless* of where $\boldsymbol{h}$ comes from, and regardless of the interior description of the system, which may very well be completely empty of matter or even contain "exotic" matter (as long as a well-posed initial value formulation exists). The EMRI problem is just a special case, where $\boldsymbol{h}$ is sourced in the background, according to the approach considered here, by a rudimentary point particle of mass m.

This opens up many interesting conceptual questions, especially with regard to the meaning of the quasilocal vacuum energy and pressure. While traditionally these have often been regarded as unphysical, to be "subtracted away" as reference terms (for the same reason that a "reference action" is often subtracted from the total gravitational action in Lagrangian formulations of GR), our analysis in this paper reveals instead that they are absolutely indispensable to accounting for self-force effects. (Indeed, the initial work [51] on the formulation of the quasilocal momentum conservation laws had similarly revealed the necessity of keeping these terms for a proper accounting of gravitational energy-momentum transfer in general.) To put it simply, the vacuum energy is what seems to play the role of the mass in the "mass times acceleration" of the self-force; the pressure term, leading to what we have called the self-pressure force, has no Newtonian analog.

Now, let us comment on our results from a more mathematical and technical point of view. When applied to a specific self-force analysis, namely, that of Gralla and Wald [33], we have been able to recover the angle average formula of Gralla [40]. The latter was put forward on the basis of a convenient mathematical argument in a Hamiltonian setting. As the quasilocal stress-energy-momentum definitions that we have been working with (namely, as given by the Brown-York tensor) recover the usual Hamiltonian definitions under appropriate conditions (stationary asymptotically flat spacetimes with a parity condition), it is reasonable that our general equation of motion [Eq. (69)]—expressing the physical flux of gravitational momentum—should thereby recover that of Gralla [Eq. (75)]—expressing an angle average in a setting where certain surface integral definitions of general-relativistic Hamiltonian notions (in particular, a Hamiltonian center of mass) can be well defined. The limitation of Gralla's equation of motion (e.g., in terms of the perturbative gauge restriction attached to it) *vis-à-vis* our general equation of motion is therefore essentially the reflection of the general limitation of Hamiltonian notions of gravitational stress-energy-momentum (as defined for a total, asymptotically flat spacetime with parity conditions) *vis-à-vis* general quasilocal notions of such concepts—of which the Hamiltonian ones arise simply as a special case.

We conjecture, although a more detailed analysis is needed to show this in detail, that the novel appearance of our self-pressure term and the precise forms of our equations of motion obtained in this paper relative to known results are attributable to our lack of imposition of a perturbative gauge, and possibly a more complicated relationship than direct identification between the simple Hamiltonian center of mass used in past self-force analyses and the choices of the centers of quasilocal frames that we have been working with here. Nevertheless, we maintain that the pressure term in general, though less intuitive than the energy term and having no Newtonian analog, ought to be viewed as just as fundamental when working with conservation laws based on the Brown-York tensor. Their intimate connection is already implied in the general identity (40), and we refer the reader to Sec. 3.3 of Ref. [51] for a detailed discussion of why the two vacuum (energy and pressure) terms are logically self-consistent. Essentially, a negative vacuum pressure acts as a positive surface tension on the quasilocal frame which is exactly compensated by the negative vacuum energy.

For carrying out practical EMRI computations, there is a manifest advantage in formulating the self-force as a closed two-surface integral around the moving particle versus standard approaches. In the latter, one typically attempts to

formulate the problem *at* the "particle location," i.e., the support of the distributional matter stress-energy-momentum tensor modeling the moving object (SCO) in the background spacetime. Of course, due to the distributional source, $\boldsymbol{h}$ actually diverges on its support, and so regularization or Green's function methods are typically employed in order to make progress. However, in principle, no such obstacles are encountered (nor are the aforementioned technical solutions needed) if the self-force is evaluated on a boundary around—very close to, but at a finite distance away from—the particle, where no formal singularity is ever encountered; $\boldsymbol{h}$ remains everywhere finite over the integration, and therefore so does the (extended) self-force functional [Eq. (11)] with it directly as its argument.

B. Outlook to future work

A numerical implementation of a concrete self-force computation using the approach developed in this paper would be arguably the most salient next step to take. To our knowledge, no numerical work has been put forth even using Gralla's angle average integral formula [40] (which would further require gauge transformations away from parity-regular gauges).

We stress here that our proposed equation of motion involving the gravitational self-force is entirely formulated and in principle valid in any choice of perturbative gauge. To our knowledge, this is the first such proposal bearing this feature. This may provide a great advantage for numerical work, as black hole metric perturbations $\boldsymbol{h}$ are often most easily computed (by solving the linearized Einstein equation, usually with a delta-function source motivated as in or similarly to the Gralla-Wald approach [33] described in Sec. IV A) in gauges that are not in the parity-regular class restricting Gralla's formula [40]. In other words, we claim that one may solve the linearized Einstein equation [Eq. (71)] for $\boldsymbol{h}^{\mathsf{X}}$ in any desired choice of gauge X, insert this $\boldsymbol{h}^{\mathsf{X}}$ into our extended GSF functional [Eq. (11)] to obtain $\mathcal{F}^{\mathsf{X}}[\boldsymbol{h}^{\mathsf{X}};\mathring{\boldsymbol{u}}^{\mathsf{X}}]$ (for some choice of background quasilocal frame with four-velocity $\mathring{\boldsymbol{u}}$), and then integrate this over a "small radius" topological two-sphere surrounding the particle (so that $\mathring{\boldsymbol{u}}$ can be approximated by the background geodesic four-velocity of the particle, $\mathring{\boldsymbol{U}}$), to obtain the full extended gravitational self-force (or correction to the motion) *directly in that gauge* X. It is easy to speculate that this may simplify some numerical issues tremendously *vis-à-vis* current approaches, where much technical machinery is needed to handle (and to do so in a sufficiently efficient way for future waveform applications) the necessary gauge transformations involving distributional source terms.

Nevertheless, further work is needed to bring the relatively abstract analysis developed in this paper into a form more readily suited for practical numerics. The most apparent technical issue to be tackled involves the fact that $\boldsymbol{h}$ is usually computed (in some kind of harmonics) in angular coordinates centered on the MBH, while the functional $\mathcal{F}[\boldsymbol{h};\mathring{\boldsymbol{u}}]$ is evaluated in angular coordinates (on a small topological two-sphere) centered on the moving particle, i.e., the SCO. A detailed understanding of the transformation between the two sets of angular coordinates is thus essential to formulate this problem numerically. This issue is discussed a bit further in Gralla's paper [40], but a detailed implementation of such a computation remains to be attempted.

The abstraction and generality of our approach may, on the other hand, also provide useful ways to address some other technical issues surrounding the self-force problem. For example, all the calculations in this paper may be carried on to second order (in the formal expansion parameter λ). This is conceptually straightforward given our basic perturbative setup, but of course requires a detailed analysis in its own right. Nonetheless, one may readily see that any higher-order correction to the motion manifestly remains here in the form of a boundary flux—only now involving nonlinear terms in the integrand. Thus, any sort of singular behavior is avoided at the level of the equations of motion in our approach, up to any order.

As another example, if ever desired (e.g., for astrophysical reasons), linear or any higher-order in r (the areal radius of the SCO boundary) effects on the correction to the motion can also be computed using our approach. Moreover, any matter fluxes (described by the usual matter stress-energy-momentum tensor, $\boldsymbol{T}$) can also be accommodated thanks to our general (gravity plus matter) conservation laws [Eq. (39)].

Furthermore, while we have applied our ideas in this paper to a specific self-force approach—that of Gralla and Wald [33]—our general formulation (Sec. III) can just as well be used in any other approach to the gravitational self-force, i.e., any other specification of a perturbative procedure (of a family of perturbed spacetimes $\{(\mathscr{M}_{(\lambda)},\boldsymbol{g}_{(\lambda)}\})$ for this problem. In other words, our approach permits any alternative specification of what is meant by a (sufficiently) "localized source" in general relativity, as our conservation expressions always involve fluxes on their boundaries and are not conditioned in any way by the exact details of their interior modeling. Thus, our equation of motion [Eq. (139)] could be used not only for a self-consistent computation (using, e.g., an approach such as that of Refs. [80,81] for solving the field equations in this context) within the Gralla-Wald approach, but also, for example, in the context of the (mathematically equivalent) self-consistent formulation of Pound [77].

Beyond the gravitational self-force, another avenue to explore from here—of interest at the very least for conceptual consistency—is how our approach handles the electromagnetic self-force problem. Although undoubtedly some conceptual parallels may be drawn between the gravitational and electromagnetic self-force problems

(see, e.g., Ref. [11]), foundationally they are usually treated as separate problems. Indeed, shortly after the paper of Gralla and Wald [33] detailing the self-force approach used in this work, Gralla *et al.* [82] put forth a similar analysis, with an analogous approach and level of rigor, of the electromagnetic self-force. It would be of great interest to apply our quasilocal conservation laws in this setting, as they can be used to account not just for gravitational but also (and in a consistent way) matter fluxes as well. It may thus prove insightful to study how the transfer of energymomentum is actually accounted for (between the gravitational and the matter sector), as in our approach we are not restricted to fixing a nondynamical metric in the spacetime. In other words, the conservation laws account completely for fluxes due to a dynamical geometry as well as matter.

ACKNOWLEDGMENTS

We thank Leor Barack and Abraham Harte for useful comments and discussions. M. O. and R. B. M. acknowledge support from the Natural Sciences and Engineering Research Council of Canada. M. O. and C. F. S. acknowledge support from Contracts No. ESP2013-47637-P, No. ESP2015-67234-P, and No. ESP2017-90084-P from the Ministry of Economy and Business of Spain (MINECO) and from Contract No. 2017-SGR-1469 from AGAUR (Catalan government). M. O. and A. D. A. M. S. acknowledge LISA CNES funding. We also acknowledge networking support by the COST Action GWverse CA16104 (Horizon 2020 Framework Programme of the European Union).

APPENDIX A: CONFORMAL KILLING VECTORS AND THE TWO-SPHERE

In this Appendix, we review some basic properties of *conformal Killing vectors* (CKVs), and in particular CKVs on the two-sphere.

A vector field $\boldsymbol{X}$ on any n-dimensional Riemannian manifold $(\mathcal{U}, \boldsymbol{g}_{\mathcal{U}}, \boldsymbol{\nabla}_{\mathcal{U}})$ is a CKV if and only if it satisfies the *conformal Killing equation*,

$$\mathcal{L}_{\boldsymbol{X}}\boldsymbol{g}_{\mathcal{U}} = \psi \boldsymbol{g}_{\mathcal{U}}, \tag{A1}$$

where $\psi \in C^{\infty}(\mathcal{U})$. This function can be determined uniquely by taking the trace of this equation, yielding

$$\psi = \frac{2}{n}\boldsymbol{\nabla}_{\mathcal{U}} \cdot \boldsymbol{X}. \tag{A2}$$

Let us now specialize to the r-radius two-sphere $(\mathbb{S}^2_r, r^2\mathfrak{S}, \mathfrak{D})$, where we denote our CKV by $\boldsymbol{\phi}$. Moreover, for ease of notation in this Appendix, the two-sphere volume form $\boldsymbol{\epsilon}_{\mathbb{S}^2}$ [Eq. (14)] is equivalently denoted by $\mathfrak{E}$, i.e., $\mathfrak{E}_{\mathfrak{ij}} = \epsilon^{\mathbb{S}^2}_{\mathfrak{ij}}$.

In this case, the conformal Killing equation (A1) is

$$\mathfrak{D}^{(\mathfrak{i}}\phi^{\mathfrak{j})} = \frac{1}{2r^2}\mathfrak{S}^{\mathfrak{ij}}\mathfrak{D}_{\mathfrak{k}}\phi^{\mathfrak{k}} \Leftrightarrow \mathfrak{D}^{\langle\mathfrak{i}}\phi^{\mathfrak{j}\rangle} = 0, \tag{A3}$$

where $\langle\cdot\cdot\rangle$ on two indices indicates taking the symmetric trace-free (STF) part. The solution to this equation can be usefully expressed in the form of a spherical harmonic decomposition in terms of the standard direction cosines of a radial unit vector in $\mathbb{R}^3$, which we denote by r^I. In spherical coordinates $\{x^{\mathfrak{i}}\} = \{\theta, \phi\}$, it is simply given by

$$r^I(\theta, \phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta). \tag{A4}$$

This satisfies the following useful identity,

$$\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} \prod_{n=1}^{\ell} r^{I_\ell} = \begin{cases} 0, & \text{for } \ell \text{ odd}, \\ \frac{4\pi}{(\ell+1)!!}\delta^{\{I_1 I_2}\cdots\delta^{I_{\ell-1}I_\ell\}}, & \text{for } \ell \text{ even}, \end{cases} \tag{A5}$$

where $(\ell+1)!! = (\ell+1)(\ell-1)\cdots 1$ and the curly brackets on the indices denote the smallest set of permutations that make the result symmetric. In particular, the $\ell = 2$ and $\ell = 4$ cases (which suffice for the calculations presented in this paper) are

$$\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^I r^J = \frac{4\pi}{3}\delta^{IJ}, \tag{A6}$$

$$\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^I r^J r^K r^L = \frac{4\pi}{15}(\delta^{IJ}\delta^{KL} + \delta^{IK}\delta^{JL} + \delta^{IL}\delta^{KJ}). \tag{A7}$$

We can construct from Eq. (A4) two sets of $\ell = 1$ spherical harmonic forms on $\mathbb{S}^2_r$, namely, the *boost generators*,

$$\mathfrak{B}^I_{\mathfrak{i}}(\theta, \phi) = \mathfrak{D}_{\mathfrak{i}} r^I = \begin{pmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\theta\sin\phi & \sin\theta\cos\phi & 0 \end{pmatrix}, \tag{A8}$$

and the *rotation generators*,

$$\mathfrak{R}^I_{\mathfrak{i}}(\theta, \phi) = \mathfrak{E}_{\mathfrak{i}}{}^{\mathfrak{j}}\mathfrak{B}^I_{\mathfrak{j}} = \epsilon^I{}_{JK} r^J \mathfrak{B}^K_{\mathfrak{i}} = \begin{pmatrix} -\sin\phi & \cos\phi & 0 \\ -\sin\theta\cos\theta\cos\phi & -\sin\theta\cos\theta\sin\phi & \sin^2\theta \end{pmatrix}, \tag{A9}$$

where ϵ_{IJK} is the volume form of $\mathbb{R}^3$. We can obtain from these the vector fields $\mathfrak{B}^{\mathfrak{i}}_I = \frac{1}{r^2}\delta_{IJ}\mathfrak{S}^{\mathfrak{ij}}\mathfrak{B}^J_{\mathfrak{j}} = \mathfrak{D}^{\mathfrak{j}} r_I$ and $\mathfrak{R}^{\mathfrak{i}}_I = \frac{1}{r^2}\delta_{IJ}\mathfrak{S}^{\mathfrak{ij}}\mathfrak{R}^J_{\mathfrak{j}} = \mathfrak{E}^{\mathfrak{i}}{}_{\mathfrak{j}}\mathfrak{B}^{\mathfrak{j}}_I$, which satisfy the Lorentz algebra

$$[\mathfrak{B}_I, \mathfrak{B}_J] = \epsilon_{IJ}{}^K \mathfrak{R}_K, \tag{A10}$$

$$[\mathfrak{R}_I, \mathfrak{B}_J] = -\epsilon_{IJ}{}^K \mathfrak{B}_K, \tag{A11}$$

$$[\mathfrak{R}_I, \mathfrak{R}_J] = -\epsilon_{IJ}{}^K \mathfrak{R}_K. \tag{A12}$$

From the above, it is possible to derive a number of useful properties:

$$r_I \mathfrak{B}^I_{\mathfrak{i}} = 0, \qquad \mathfrak{D}_{\mathfrak{i}} \mathfrak{B}^I_{\mathfrak{j}} = -\mathfrak{S}_{\mathfrak{ij}} r^I \Rightarrow \mathfrak{S}^{\mathfrak{ij}} \mathfrak{D}_{\mathfrak{i}} \mathfrak{B}^I_{\mathfrak{j}} = -2r^I, \tag{A13}$$

$$r_I \mathfrak{R}^I_{\mathfrak{i}} = 0, \qquad \mathfrak{D}_{\mathfrak{i}} \mathfrak{R}^I_{\mathfrak{j}} = \mathfrak{E}_{\mathfrak{ij}} r^I \Rightarrow \mathfrak{S}^{\mathfrak{ij}} \mathfrak{D}_{\mathfrak{i}} \mathfrak{R}^I_{\mathfrak{j}} = 0. \tag{A14}$$

Using these, one can show that the sets of $\ell = 1$ vector fields $\mathfrak{B}_I$ and $\mathfrak{R}_I$ all satisfy the conformal Killing equation, i.e.,

$$\mathfrak{D}^{\langle \mathfrak{i}} \mathfrak{B}^{\mathfrak{j}\rangle}_I = 0 = \mathfrak{D}^{\langle \mathfrak{i}} \mathfrak{R}^{\mathfrak{j}\rangle}_I. \tag{A15}$$

Finally, we give a list of useful relations for various contractions involving these vector fields,

$$\mathfrak{S}^{\mathfrak{ij}} \mathfrak{B}^I_{\mathfrak{i}} \mathfrak{B}^J_{\mathfrak{j}} = \mathfrak{S}^{\mathfrak{ij}} \mathfrak{R}^I_{\mathfrak{i}} \mathfrak{R}^J_{\mathfrak{j}} = -\mathfrak{E}^{\mathfrak{ij}} \mathfrak{B}^I_{\mathfrak{i}} \mathfrak{R}^J_{\mathfrak{j}} = P^{IJ}, \tag{A16}$$

$$\mathfrak{S}^{\mathfrak{ij}} \mathfrak{B}^I_{\mathfrak{i}} \mathfrak{R}^J_{\mathfrak{j}} = \mathfrak{E}_{\mathfrak{ij}} \mathfrak{B}^{\mathfrak{i}}_I \mathfrak{B}^{\mathfrak{j}}_J = \mathfrak{E}_{\mathfrak{ij}} \mathfrak{R}^{\mathfrak{i}}_I \mathfrak{R}^{\mathfrak{j}}_J = \epsilon^{IJK} r_K, \tag{A17}$$

$$\delta_{IJ} \mathfrak{B}^I_{\mathfrak{i}} \mathfrak{B}^J_{\mathfrak{j}} = \delta_{IJ} \mathfrak{R}^I_{\mathfrak{i}} \mathfrak{R}^J_{\mathfrak{j}} = \mathfrak{S}_{\mathfrak{ij}}, \tag{A18}$$

$$\delta_{IJ} \mathfrak{B}^I_{\mathfrak{i}} \mathfrak{R}^J_{\mathfrak{j}} = -\mathfrak{E}_{\mathfrak{ij}}, \tag{A19}$$

$$\epsilon_{IJK} \mathfrak{B}^I_{\mathfrak{i}} \mathfrak{B}^J_{\mathfrak{j}} = \epsilon_{IJK} \mathfrak{R}^I_{\mathfrak{i}} \mathfrak{R}^J_{\mathfrak{j}} = \mathfrak{E}_{\mathfrak{ij}} r_K, \tag{A20}$$

$$\epsilon_{IJK} \mathfrak{B}^I_{\mathfrak{i}} \mathfrak{R}^J_{\mathfrak{j}} = \mathfrak{S}_{\mathfrak{ij}} r_K, \tag{A21}$$

where $P^{IJ} = \delta^{IJ} - r^I r^J$ projects vectors perpendicular to the radial direction.

Now, we have everything in hand to formulate the general solution to the conformal Killing equation (A3) on $\mathbb{S}^2_r$; it can be expanded as

$$\phi^{\mathfrak{j}} = \frac{1}{r}\left[\mathfrak{D}^{\mathfrak{j}}\left(\sum_{\ell\in\mathbb{N}} \Phi^{I_1\cdots I_\ell} \prod_{n=1}^{\ell} r_{I_n}\right) + \mathfrak{E}^{\mathfrak{j}}{}_{\mathfrak{k}} \mathfrak{D}^{\mathfrak{k}}\left(\sum_{\ell\in\mathbb{N}} \Psi^{I_1\cdots I_\ell} \prod_{n=1}^{\ell} r_{I_n}\right)\right], \tag{A22}$$

$$= \frac{1}{r}\left[(\Phi^I \mathfrak{B}^{\mathfrak{j}}_I + \Psi^I \mathfrak{R}^{\mathfrak{j}}_I) + \sum_{\ell\geq 2} \ell (\Phi^{I_1\cdots I_\ell} \mathfrak{B}^{\mathfrak{j}}_{I_1} + \Psi^{I_1\cdots I_\ell} \mathfrak{R}^{\mathfrak{j}}_{I_1}) \prod_{n=2}^{\ell} r_{I_n}\right], \tag{A23}$$

where to write the second equality we have used the fact that $\Phi^{I_1\cdots I_\ell}$ and $\Psi^{I_1\cdots I_\ell}$ are symmetric in their indices. One can thus check that

$$\mathfrak{D}^{\mathfrak{i}}\phi^{\mathfrak{j}} = \chi\Bigg[(-\Phi^I \mathfrak{S}^{\mathfrak{ij}} + \Psi^I \mathfrak{E}^{\mathfrak{ij}}) r_I + 2(\Phi^{IJ}\{-\mathfrak{S}^{\mathfrak{ij}} r_I r_J + \mathfrak{B}^{\mathfrak{i}}_I \mathfrak{B}^{\mathfrak{j}}_J\} + \Psi^{IJ}\{\mathfrak{E}^{\mathfrak{ij}} r_I r_J + \mathfrak{B}^{\mathfrak{i}}_I \mathfrak{R}^{\mathfrak{j}}_J\}) + \sum_{\ell\geq 3} \ell (\Phi^{I_1\cdots I_\ell}\{-\mathfrak{S}^{\mathfrak{ij}} r_{I_1} r_{I_2} + (\ell-1)\mathfrak{B}^{\mathfrak{i}}_{I_1} \mathfrak{B}^{\mathfrak{j}}_{I_2}\} + \Psi^{I_1\cdots I_\ell}\{\mathfrak{E}^{\mathfrak{ij}} r_{I_1} r_{I_2} + (\ell-1)\mathfrak{B}^{\mathfrak{i}}_{I_1} \mathfrak{R}^{\mathfrak{j}}_{I_2}\}) \prod_{n=2}^{\ell} r_{I_n}\Bigg]. \tag{A24}$$

We are interested in working with the $\ell = 1$ and $\ell = 2$ parts of $\boldsymbol{\phi}$ corresponding to linear momentum only ($\boldsymbol{\Psi} = 0$):

$$\phi^{\mathfrak{i}}_{\ell=1} = \frac{1}{r} \Phi^I \mathfrak{B}^{\mathfrak{i}}_I, \tag{A25}$$

$$\phi^{\mathfrak{i}}_{\ell=2} = \frac{2}{r} \Phi^{IJ} \mathfrak{B}^{\mathfrak{i}}_I r_J. \tag{A26}$$

APPENDIX B: DETAILED CALCULATIONS FOR THE EQUATIONS OF MOTION

1. Rigid quasilocal frame inertial with the background point particle

For the background momentum change, the $\ell = 1$ and $\ell = 2$ parts are computed, respectively, as follows:

$$(\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})})_{(0)} = -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2 \mathring{N}(\mathring{\mathcal{E}}\mathring{\boldsymbol{\alpha}}\cdot\boldsymbol{\phi}_{\ell=1} + \mathring{\mathrm{P}}\mathfrak{D}\cdot\boldsymbol{\phi}_{\ell=1}) \tag{B1}$$

$$= -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2 [1 + \mathcal{O}(r^2)]\left(\left[-\frac{2}{\kappa r} + \mathcal{O}(r)\right]\frac{1}{r^2}\mathfrak{S}^{\mathfrak{ij}}[r^2 \mathring{E}_{JK} \mathfrak{B}^J_{\mathfrak{i}} r^K + \mathcal{O}(r^3)][r\Phi_I \mathfrak{B}^I_{\mathfrak{j}}]\right. \tag{B2}$$

$$+\left[-\frac{1}{\kappa r}+\mathcal{O}(r)\right]\frac{1}{r^2}\mathfrak{S}^{\mathfrak{ij}}\mathfrak{D}_{\mathfrak{i}}[r\Phi_I\mathfrak{B}^I_{\mathfrak{j}}]\Big) \tag{B3}$$

$$=\frac{1}{\kappa}\Phi_I\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}[1+\mathcal{O}(r^2)]([2+\mathcal{O}(r^2)][\mathcal{O}(r^2)]+[1+\mathcal{O}(r^2)]\mathfrak{S}^{\mathfrak{ij}}\mathfrak{D}_{\mathfrak{i}}\mathfrak{B}^I_{\mathfrak{j}}) \tag{B4}$$

$$=\frac{1}{\kappa}\Phi_I\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}(-2r^I)+\mathcal{O}(r^2) \tag{B5}$$

$$=\mathcal{O}(r^2), \tag{B6}$$

and

$$(\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=2})})_{(0)}=-\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}r^2\mathring{N}(\mathring{\mathcal{E}}\mathring{\boldsymbol{\alpha}}\cdot\boldsymbol{\phi}_{\ell=2}+\mathring{\mathrm{P}}\mathfrak{D}\cdot\boldsymbol{\phi}_{\ell=2}) \tag{B7}$$

$$=-\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}r^2[1+\mathcal{O}(r^2)]\left(\left[-\frac{2}{\kappa r}+\mathcal{O}(r)\right]\frac{1}{r^2}\mathfrak{S}^{\mathfrak{ij}}[r^2\mathring{E}_{JK}\mathfrak{B}^J_{\mathfrak{i}}r^K+\mathcal{O}(r^3)][2r\Phi_{IJ}r^I\mathfrak{B}^J_{\mathfrak{j}}] \tag{B8}$$

$$+\left[-\frac{1}{\kappa r}+\mathcal{O}(r)\right]\frac{1}{r^2}\mathfrak{S}^{\mathfrak{ij}}\mathfrak{D}_{\mathfrak{i}}[2r\Phi_{IJ}r^I\mathfrak{B}^J_{\mathfrak{j}}]\Big) \tag{B9}$$

$$=\frac{1}{\kappa}\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}[1+\mathcal{O}(r^2)]([2+\mathcal{O}(r^2)][\mathcal{O}(r^2)]+[1+\mathcal{O}(r^2)]\mathfrak{S}^{\mathfrak{ij}}\mathfrak{D}_{\mathfrak{i}}[2\Phi_{IJ}r^I\mathfrak{B}^J_{\mathfrak{j}}]) \tag{B10}$$

$$=\frac{2}{\kappa}\Phi_{IJ}\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}(\mathfrak{S}^{\mathfrak{ij}}\mathfrak{D}_{\mathfrak{i}}[r^I\mathfrak{B}^J_{\mathfrak{j}}])+\mathcal{O}(r^2) \tag{B11}$$

$$=\frac{2}{\kappa}\Phi_{IJ}\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}(\mathfrak{S}^{\mathfrak{ij}}\mathfrak{B}^I_{\mathfrak{i}}\mathfrak{B}^J_{\mathfrak{j}}+r^I\mathfrak{S}^{\mathfrak{ij}}\mathfrak{D}_{\mathfrak{i}}\mathfrak{B}^J_{\mathfrak{j}})+\mathcal{O}(r^2) \tag{B12}$$

$$=\frac{2}{\kappa}\Phi_{IJ}\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}(P^{IJ}-2r^Ir^J)+\mathcal{O}(r^2) \tag{B13}$$

$$=\mathcal{O}(r^2). \tag{B14}$$

For the perturbed momentum change, we have

$$\delta\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta N)}=-\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}r^2\delta N(\mathring{\mathcal{E}}\mathring{\boldsymbol{\alpha}}\cdot\boldsymbol{\phi}_{\ell=1}+\mathring{\mathrm{P}}\mathfrak{D}\cdot\boldsymbol{\phi}_{\ell=1}) \tag{B15}$$

$$=-\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}r^2\delta N\left(\left[-\frac{2}{\kappa r}+\mathcal{O}(r)\right]\frac{1}{r^2}\mathfrak{S}^{\mathfrak{ij}}[r^2\mathring{E}_{JK}\mathfrak{B}^J_{\mathfrak{i}}r^K+\mathcal{O}(r^3)][r\Phi_I\mathfrak{B}^I_{\mathfrak{j}}] \tag{B16}$$

$$+\left[-\frac{1}{\kappa r}+\mathcal{O}(r)\right]\frac{1}{r^2}\mathfrak{S}^{\mathfrak{ij}}\mathfrak{D}_{\mathfrak{i}}[r\Phi_I\mathfrak{B}^I_{\mathfrak{j}}]\Big) \tag{B17}$$

$$=\frac{1}{\kappa}\Phi_I\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}\delta N([2+\mathcal{O}(r^2)][r^2\mathring{E}_{JK}r^KP^{JI}+\mathcal{O}(r^3)]+[1+\mathcal{O}(r^2)][-2r^I]) \tag{B18}$$

$$=-\frac{2}{\kappa}\Phi_I\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}\delta Nr^I+\mathcal{O}(r^2), \tag{B19}$$

$$\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\mathcal{E})} = -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2 \mathring{N}\delta\mathcal{E}\mathring{\boldsymbol{\alpha}}\cdot\boldsymbol{\phi} \tag{B20}$$

$$= -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2[1+\mathcal{O}(r^2)]\left[\frac{m}{4\pi r^2}\right]\frac{1}{r^2}\mathfrak{S}^{\mathrm{ij}}[r^2\mathring{E}_{JK}\mathfrak{B}^J_{\mathrm{i}}r^K+\mathcal{O}(r^3)][r\Phi_I\mathfrak{B}^I_{\mathrm{j}}] \tag{B21}$$

$$= -\frac{m}{4\pi}r\Phi_I\mathring{E}_{JK}\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}[1+\mathcal{O}(r^2)][r^K P^{JI}]+\mathcal{O}(r^2) \tag{B22}$$

$$= \mathcal{O}(r^2), \tag{B23}$$

and

$$\delta\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\mathrm{P})} = -\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2} r^2\mathring{N}\delta\mathrm{P}\mathfrak{D}\cdot\boldsymbol{\phi}_{\ell=1} \tag{B24}$$

$$= -\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2} r^2\left[1+\frac{1}{2}r^2\mathring{E}_{IJ}r^Ir^J+\mathcal{O}(r^3)\right]\left[\frac{m}{8\pi r^2}-\frac{1}{\kappa}\delta a_{\boldsymbol{n}}\right]\frac{1}{r^2}\mathfrak{S}^{\mathrm{ij}}\mathfrak{D}_{\mathrm{i}}[r\Phi_I\mathfrak{B}^I_{\mathrm{j}}] \tag{B25}$$

$$= -r\Phi_I\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}\left[1+\frac{1}{2}r^2\mathring{E}_{IJ}r^Ir^J+\mathcal{O}(r^3)\right]\left[\frac{m}{8\pi r^2}-\frac{1}{\kappa}\delta a_{\boldsymbol{n}}\right](-2r^I) \tag{B26}$$

$$= \mathcal{O}(r^2). \tag{B27}$$

2. Rigid quasilocal frame inertial with the moving object in the perturbed spacetime

The $\ell=1$ background momentum change is given by $(\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})})_{(0)} = (\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{\mathring{(\mathcal{E})}})_{(0)} + (\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{\mathring{(\mathrm{P})}})_{(0)}$. We have

$$(\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{\mathring{(\mathcal{E})}})_{(0)} = -\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2} r^2\mathring{N}\,\mathring{\mathcal{E}}\,\mathring{\boldsymbol{\alpha}}\cdot\boldsymbol{\phi}_{\ell=1} \tag{B28}$$

$$= -\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2} r^2[1+\mathcal{O}(r)]\left[-\frac{2}{\kappa r}+\mathcal{O}(r)\right]\frac{1}{r^2}\mathfrak{S}^{\mathrm{ij}}[rA_K\mathfrak{B}^K_{\mathrm{i}}+\mathcal{O}(r^2)][r\Phi_I\mathfrak{B}^I_{\mathrm{j}}] \tag{B29}$$

$$= \frac{2}{\kappa}\Phi_I\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}[1+\mathcal{O}(r)][1+\mathcal{O}(r^2)]\mathfrak{S}^{\mathrm{ij}}[rA_J\mathfrak{B}^J_{\mathrm{i}}\mathfrak{B}^I_{\mathrm{j}}+\mathcal{O}(r^2)] \tag{B30}$$

$$= \frac{2r}{\kappa}\Phi_IA_J\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}\mathfrak{S}^{\mathrm{ij}}\mathfrak{B}^J_{\mathrm{i}}\mathfrak{B}^I_{\mathrm{j}}+\mathcal{O}(r^2) \tag{B31}$$

$$= \frac{2r}{\kappa}\Phi_IA_J\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}P^{IJ}+\mathcal{O}(r^2) \tag{B32}$$

$$= \frac{16\pi}{3\kappa}r\Phi^IA_I+\mathcal{O}(r^2), \tag{B33}$$

using the fact that $\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}P_{IJ} = \int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2}(\delta_{IJ}-r_Ir_J) = 4\pi(1-\frac{1}{3})\delta_{IJ} = \frac{8\pi}{3}\delta_{IJ}$, and

$$(\dot{\mathrm{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{\mathring{(\mathrm{P})}})_{(0)} = -\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2} r^2\mathring{N}\,\mathring{\mathrm{P}}\,\mathfrak{D}\cdot\boldsymbol{\phi}_{\ell=1} \tag{B34}$$

$$= -\int_{\mathbb{S}^2_r}\boldsymbol{\epsilon}_{\mathbb{S}^2} r^2[1+rA_Jr^J+\mathcal{O}(r^2)]\left[-\frac{1}{\kappa r}-\frac{1}{\kappa}A_Kr^K+\mathcal{O}(r)\right]\frac{1}{r^2}\mathfrak{S}^{\mathrm{ij}}\mathfrak{D}_{\mathrm{i}}[r\Phi_I\mathfrak{B}^I_{\mathrm{j}}] \tag{B35}$$

$$= \frac{1}{\kappa}\Phi_I \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2}[1 + rA_J r^J + \mathcal{O}(r^2)][1 + rA_K r^K + \mathcal{O}(r^2)]\mathfrak{S}^{\mathfrak{ij}}\mathfrak{D}_{\mathfrak{i}}\mathfrak{B}^I_{\mathfrak{j}} \tag{B36}$$

$$= -\frac{2}{\kappa}\Phi_I \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2}[1 + rA_J r^J + \mathcal{O}(r^2)][1 + rA_K r^K + \mathcal{O}(r^2)]r^I \tag{B37}$$

$$= -\frac{4}{\kappa} r\Phi_I A_J \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^J r^I + \mathcal{O}(r^2) \tag{B38}$$

$$= -\frac{16\pi}{3\kappa} r\Phi^I A_I + \mathcal{O}(r^2), \tag{B39}$$

where in the fifth equality the fact that $\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r_I = 0$ leads to the vanishing of the $\mathcal{O}(r^0)$ term. Adding the two results, $(\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})})_{(0)} = \mathcal{O}(r^2)$. For the $\ell = 2$ part, we similarly have $(\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=2})})_{(0)} = (\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=2})}_{(\mathring{\mathcal{E}})})_{(0)} + (\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=2})}_{(\mathring{\mathrm{P}})})_{(0)}$. First,

$$(\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=2})}_{(\mathring{\mathcal{E}})})_{(0)} = -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2 \mathring{N}\mathring{\mathcal{E}}\mathring{\boldsymbol{\alpha}}\cdot\boldsymbol{\phi}_{\ell=2} \tag{B40}$$

$$= -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2[1 + rA_N r^N + \mathcal{O}(r^2)]\left[-\frac{2}{\kappa r} + \mathcal{O}(r)\right] \tag{B41}$$

$$\times \frac{1}{r^2}\mathfrak{S}^{\mathfrak{ij}}[rA_K\mathfrak{B}^K_{\mathfrak{i}} + r^2(\mathring{E}_{LM} - A_L A_M)\mathfrak{B}^L_{\mathfrak{i}} r^M + \mathcal{O}(r^3)][2r\Phi_{IJ} r^I \mathfrak{B}^J_{\mathfrak{j}}] \tag{B42}$$

$$= \frac{4}{\kappa} r\Phi_{IJ} \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2}[1 + rA_N r^N + \mathcal{O}(r^2)][1 + \mathcal{O}(r^2)] \tag{B43}$$

$$\times \mathfrak{S}^{\mathfrak{ij}}[A_K\mathfrak{B}^K_{\mathfrak{i}} + r(\mathring{E}_{LM} - A_L A_M)\mathfrak{B}^L_{\mathfrak{i}} r^M + \mathcal{O}(r^2)][r^I \mathfrak{B}^J_{\mathfrak{j}}] \tag{B44}$$

$$= \frac{4}{\kappa} r^2\Phi_{IJ} \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^I \mathfrak{S}^{\mathfrak{ij}}\mathfrak{B}^J_{\mathfrak{j}}[(\mathring{E}_{LM} - A_L A_M)\mathfrak{B}^L_{\mathfrak{i}} r^M + A_N r^N A_K \mathfrak{B}^K_{\mathfrak{i}}] + \mathcal{O}(r^3) \tag{B45}$$

$$= \frac{4}{\kappa} r^2\Phi_{IJ}\mathring{E}_{LM} \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^I \mathfrak{S}^{\mathfrak{ij}}\mathfrak{B}^J_{\mathfrak{j}}\mathfrak{B}^L_{\mathfrak{i}} r^M + \mathcal{O}(r^3) \tag{B46}$$

$$= \frac{4}{\kappa} r^2\Phi_{IJ}\mathring{E}_{LM} \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} P^{LJ} r^I r^M + \mathcal{O}(r^3) \tag{B47}$$

$$= \frac{16\pi}{5\kappa} r^2\Phi^{IJ}\mathring{E}_{IJ} + \mathcal{O}(r^3), \tag{B48}$$

where in the fourth equality the $\mathcal{O}(r)$ term vanishes upon integration since it is an $\ell = 3$ spherical harmonic, and in the seventh we have used the fact that

$$\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} P^{LJ} r^I r^M = \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2}(\delta^{LJ} - r^L r^J) r^I r^M = \frac{4\pi}{15}(4\delta^{LJ}\delta^{IM} - \delta^{LI}\delta^{JM} - \delta^{LM}\delta^{JI}), \tag{B49}$$

along with the fact that Φ_{IJ} and $\mathring{E}_{IJ}$ are STF terms. Next,

$$(\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=2})}_{(\mathring{\mathrm{P}})})_{(0)} = -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2 \mathring{N}\mathring{\mathrm{P}}\mathfrak{D}\cdot\boldsymbol{\phi}_{\ell=2} \tag{B50}$$

$$= -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2[1 + rA_L r^L + \mathcal{O}(r^2)]\left[-\frac{1}{\kappa r} - \frac{1}{\kappa}A_K r^K + \mathcal{O}(r)\right]\frac{1}{r^2}\mathfrak{S}^{\mathfrak{ij}}\mathfrak{D}_{\mathfrak{i}}[2r\Phi_{IJ} r^I \mathfrak{B}^J_{\mathfrak{j}}] \tag{B51}$$

$$= \frac{2}{\kappa}\Phi_{IJ}\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2}[1 + rA_L r^L + \mathcal{O}(r^2)][1 + rA_K r^K + \mathcal{O}(r^2)]\mathfrak{S}^{\mathfrak{ij}}[\mathfrak{B}^I_{\mathfrak{i}}\mathfrak{B}^J_{\mathfrak{j}} + r^I\mathfrak{D}_{\mathfrak{i}}\mathfrak{B}^J_{\mathfrak{j}}] \tag{B52}$$

$$= \frac{2}{\kappa}\Phi_{IJ}\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2}[1 + rA_L r^L + \mathcal{O}(r^2)][1 + rA_K r^K + \mathcal{O}(r^2)][P^{IJ} + r^I(-2r^J)] \tag{B53}$$

$$= \frac{2}{\kappa}\Phi_{IJ}\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2}[1 + rA_L r^L + \mathcal{O}(r^2)][1 + rA_K r^K + \mathcal{O}(r^2)][P^{IJ} - 2r^I r^J] \tag{B54}$$

$$= \mathcal{O}(r^2). \tag{B55}$$

Here, we would need the next (linear in r) term in the expansion of $\mathring{\mathrm{P}}$ (the one for $\mathring{N}$ is easy to obtain) in order to explicitly calculate the $\mathcal{O}(r^2)$ term above. In any case, we have found, just as in the PP-inertial case, $(\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=2})})_{(0)} = \mathcal{O}(r^2)$.

For the $\mathcal{O}(\lambda)$ part, as in the point-particle-inertial case, we will have

$$\delta\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})} = \sum_{Q\in\{\delta N,\delta\mathcal{E},\delta\boldsymbol{\alpha},\delta\mathrm{P},\delta\boldsymbol{D}\}} \delta\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(Q)}. \tag{B56}$$

We compute these terms one by one. First,

$$\delta\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta N)} = -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2\delta N(\mathring{\mathcal{E}}\,\mathring{\boldsymbol{\alpha}}\cdot\boldsymbol{\phi}_{\ell=1} + \mathring{\mathrm{P}}\mathfrak{D}\cdot\boldsymbol{\phi}_{\ell=1}) \tag{B57}$$

$$= -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2\delta N\left(\left[-\frac{2}{\kappa r} + \mathcal{O}(r)\right]\frac{1}{r^2}\mathfrak{S}^{\mathfrak{ij}}[rA_K\mathfrak{B}^K_{\mathfrak{i}} + \mathcal{O}(r^2)][r\Phi_I\mathfrak{B}^I_{\mathfrak{j}}]\right. \tag{B58}$$

$$\left. + \left[-\frac{1}{\kappa r} + \mathcal{O}(1)\right]\frac{1}{r^2}\mathfrak{S}^{\mathfrak{ij}}\mathfrak{D}_{\mathfrak{i}}[r\Phi_I\mathfrak{B}^I_{\mathfrak{j}}]\right) \tag{B59}$$

$$= \frac{1}{\kappa}\Phi_I\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2}\delta N([2 + \mathcal{O}(r^2)]\mathfrak{S}^{\mathfrak{ij}}[rA_K\mathfrak{B}^K_{\mathfrak{i}}\mathfrak{B}^I_{\mathfrak{j}} + \mathcal{O}(r^2)] + [1 + \mathcal{O}(r)]\mathfrak{S}^{\mathfrak{ij}}\mathfrak{D}_{\mathfrak{i}}\mathfrak{B}^I_{\mathfrak{j}}) \tag{B60}$$

$$= -\frac{2}{\kappa}\Phi_I\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2}\delta N r^I + \mathcal{O}(r). \tag{B61}$$

Next, we have

$$\delta\dot{\mathfrak{p}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\mathcal{E})} = -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2\mathring{N}\delta\mathcal{E}\mathring{\boldsymbol{\alpha}}\cdot\boldsymbol{\phi} \tag{B62}$$

$$= -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2[1 + \mathcal{O}(r)]\left[\frac{m}{4\pi r^2}\right]\frac{1}{r^2}\mathfrak{S}^{\mathfrak{ij}}[rA_K\mathfrak{B}^K_{\mathfrak{i}} + \mathcal{O}(r^2)][r\Phi_I\mathfrak{B}^I_{\mathfrak{j}}] \tag{B63}$$

$$= -\frac{m}{4\pi}\Phi_I A_K\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} P^{KI} + \mathcal{O}(r) \tag{B64}$$

$$= -\frac{m}{4\pi}\Phi_I A_K\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} P^{KI} + \mathcal{O}(r) \tag{B65}$$

$$= -\frac{2}{3}m\Phi_I A^I + \mathcal{O}(r), \tag{B66}$$

and

$$\delta\dot{\mathrm{P}}^{(\boldsymbol{\phi}_{\ell=1})}_{(\delta\mathrm{P})} = -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2 \mathring{N} \delta\mathrm{P} \boldsymbol{\mathfrak{D}} \cdot \boldsymbol{\phi}_{\ell=1} \tag{B67}$$

$$= -\int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^2 [1 + rA_L r^L + \mathcal{O}(r^2)] \left[\frac{m}{8\pi r^2} - \frac{1}{\kappa}\delta a_{\boldsymbol{n}}\right] \frac{1}{r^2} \mathfrak{S}^{\mathrm{ij}} \mathfrak{D}_{\mathrm{i}} [r\Phi_I \mathfrak{B}^I_{\mathrm{j}}] \tag{B68}$$

$$= -\Phi_I \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} [1 + rA_L r^L] \left[\frac{m}{8\pi r^2}\right] r(-2r^I) + \mathcal{O}(r) \tag{B69}$$

$$= \frac{m}{4\pi} \Phi_I A_L \int_{\mathbb{S}^2_r} \boldsymbol{\epsilon}_{\mathbb{S}^2} r^L r^I + \mathcal{O}(r) \tag{B70}$$

$$= \frac{1}{3} m \Phi_I A^I + \mathcal{O}(r). \tag{B71}$$

[1] B. P. Abbott *et al.*, GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs, Phys. Rev. X **9**, 031040 (2019).

[2] M. Armano *et al.*, Sub-Femto-g Free Fall for Space-Based Gravitational Wave Observatories: LISA Pathfinder Results, Phys. Rev. Lett. **116**, 231101 (2016).

[3] M. Armano *et al.*, Beyond the Required LISA Free-Fall Performance: New LISA Pathfinder Results Down to 20 muHz, Phys. Rev. Lett. **120**, 061101 (2018).

[4] P. Amaro-Seoane *et al.*, The gravitational Universe, arXiv:1305.5720.

[5] P. Amaro-Seoane *et al.*, Laser Interferometer Space Antenna, arXiv:1702.00786.

[6] M. Colpi and A. Sesana, Gravitational wave sources in the era of multi-band gravitational wave astronomy, in *An Overview of Gravitational Waves: Theory, Sources and Detection*, edited by A. Gerard and E. Plagnol (World Scientific, Singapore, 2016), p. 43.

[7] M. Celoria, R. Oliveri, A. Sesana, and M. Mapelli, Lecture notes on black hole binary astrophysics, arXiv:1807.11489.

[8] P. Amaro-Seoane, X-MRIs: Extremely large mass-ratio inspirals, Phys. Rev. D **99**, 123025 (2019).

[9] P. Amaro-Seoane, Relativistic dynamics and extreme mass ratio inspirals, Living Rev. Relativity **21**, 4 (2018).

[10] S. Babak, J. Gair, A. Sesana, E. Barausse, C. F. Sopuerta, C. P. Berry, E. Berti, P. Amaro-Seoane, A. Petiteau, and A. Klein, Science with the space-based interferometer LISA. V. Extreme mass-ratio inspirals, Phys. Rev. D **95**, 103012 (2017).

[11] L. Barack and A. Pound, Self-force and radiation reaction in general relativity, Rep. Prog. Phys. **82**, 016904 (2019).

[12] C. P. L. Berry, S. A. Hughes, C. F. Sopuerta, A. J. K. Chua, A. Heffernan, K. Holley-Bockelmann, D. P. Mihaylov, M. C. Miller, and A. Sesana, The unique potential of extreme mass-ratio inspirals for gravitational-wave astronomy, arXiv:1903.03686.

[13] L. Schwartz, *Théorie des Distributions* (Hermann, Paris, 1957).

[14] L. Schwartz, Sur l'impossibilité de la multiplication des distributions, C.R. Hebd. Seances Acad. Sci. **29**, 847 (1954).

[15] C. K. Li, A review on the products of distributions, in *Mathematical Methods in Engineering* edited by K. Tas, J. A. Tenreiro Machado, and D. Baleanu (Springer, Dordrecht, 2007), p. 71.

[16] J. F. Colombeau, Nonlinear generalized functions: Their origin, some developments and recent advances, São Paulo J. Math. Sci. **7**, 201 (2013), https://www.revistas.usp.br/spjournal/article/view/88606.

[17] E. Bottazzi, Grid functions of nonstandard analysis in the theory of distributions and in partial differential equations, Adv. Math. **345**, 429 (2019).

[18] A. Gsponer, The classical point electron in Colombeau's theory of nonlinear generalized functions, J. Math. Phys. (N.Y.) **49**, 102901 (2008).

[19] R. Steinbauer and J. A. Vickers, The use of generalized functions and distributions in general relativity, Classical Quantum Gravity **23**, R91 (2006).

[20] R. Geroch and J. Traschen, Strings and other distributional sources in general relativity, Phys. Rev. D **36**, 1017 (1987).

[21] L. Barack and D. A. Golbourn, Scalar-field perturbations from a particle orbiting a black hole using numerical evolution in 2+1 dimensions, Phys. Rev. D **76**, 044020 (2007).

[22] L. Barack, D. A. Golbourn, and N. Sago, m-mode regularization scheme for the self-force in Kerr spacetime, Phys. Rev. D **76**, 124036 (2007).

[23] I. Vega and S. Detweiler, Regularization of fields for s elf-force problems in curved spacetime: Foundations and a time-domain application, Phys. Rev. D **77**, 084008 (2008).

[24] M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower, Accurate Evolutions of Orbiting Black-Hole

Binaries without Excision, Phys. Rev. Lett. **96**, 111101 (2006).
[25] J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. van Meter, Gravitational-Wave Extraction from an Inspiraling Configuration of Merging Black Holes, Phys. Rev. Lett. **96**, 111102 (2006).
[26] A. Pound, Second-order gravitational self-force, Phys. Rev. Lett. **109**, 051101 (2012).
[27] S. E. Gralla, Second-order gravitational self-force, Phys. Rev. D **85**, 124011 (2012).
[28] A. Pound, Nonlinear gravitational self-force: Second-order equation of motion, Phys. Rev. D **95**, 104056 (2017).
[29] A. Einstein, L. Infeld, and B. Hoffmann, The gravitational equations and the problem of motion, Ann. Math. **39**, 65 (1938).
[30] R. Geroch and P. S. Jang, Motion of a body in general relativity, J. Math. Phys. (N.Y.) **16**, 65 (1975).
[31] J. Ehlers and R. Geroch, Equation of motion of small bodies in relativity, Ann. Phys. (N.Y.) **309**, 232 (2004).
[32] J. O. Weatherall, Geometry and motion in general relativity, arXiv:1810.09046.
[33] S. E. Gralla and R. M. Wald, A rigorous derivation of gravitational self-force, Classical Quantum Gravity **25**, 205009 (2008).
[34] Y. Mino, M. Sasaki, and T. Tanaka, Gravitational radiation reaction to a particle motion, Phys. Rev. D **55**, 3457 (1997).
[35] T. C. Quinn and R. M. Wald, Axiomatic approach to electromagnetic and gravitational radiation reaction of particles in curved spacetime, Phys. Rev. D **56**, 3381 (1997).
[36] E. Poisson, A. Pound, and I. Vega, The motion of point particles in curved spacetime, Living Rev. Relativity **14**, 7 (2011).
[37] A. D. A. M. Spallicci, P. Ritter, and S. Aoudia, Self-force driven motion in curved spacetime, Int. J. Geom. Methods Mod. Phys. **11**, 1450072 (2014).
[38] S. Detweiler and B. F. Whiting, Self-force via a Green's function decomposition, Phys. Rev. D **67**, 024025 (2003).
[39] L. Barack and C. O. Lousto, Perturbations of Schwarzschild black holes in the Lorenz gauge: Formulation and numerical implementation, Phys. Rev. D **72**, 104026 (2005).
[40] S. E. Gralla, Gauge and averaging in gravitational self-force, Phys. Rev. D **84**, 084050 (2011).
[41] T. Regge and C. Teitelboim, Role of surface integrals in the Hamiltonian formulation of general relativity, Ann. Phys. **88**, 286 (1974).
[42] A. Pound, C. Merlin, and L. Barack, Gravitational self-force from radiation-gauge metric perturbations, Phys. Rev. D **89**, 024009 (2014).
[43] P. A. M. Dirac, Classical theory of radiating electrons, Proc. R. Soc. A **167**, 148 (1938).
[44] B. S. DeWitt and R. W. Brehme, Radiation damping in a gravitational field, Ann. Phys. (N.Y.) **9**, 220 (1960).
[45] E. Poisson, An introduction to the Lorentz-Dirac equation, arXiv:gr-qc/9912045.
[46] L. M. Burko, Self-force on static charges in Schwarzschild spacetime, Classical Quantum Gravity **17**, 227 (2000).
[47] A. I. Harte, Self-forces from generalized Killing fields, Classical Quantum Gravity **25**, 235020 (2008).
[48] A. I. Harte, Electromagnetic self-forces and generalized Killing fields, Classical Quantum Gravity **26**, 155015 (2009).
[49] A. I. Harte, Mechanics of extended masses in general relativity, Classical Quantum Gravity **29**, 055012 (2012).
[50] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[51] R. J. Epp, P. L. McGrath, and R. B. Mann, Momentum in general relativity: Local versus quasilocal conservation laws, Classical Quantum Gravity **30**, 195019 (2013).
[52] L. B. Szabados, Quasi-local energy-momentum and angular momentum in GR: A review article, Living Rev. Relativity **7**, 4 (2004).
[53] J. L. Jaramillo and E. Gourgoulhon, Mass and angular momentum in general relativity, in *Mass and Motion in General Relativity*, edited by L. Blanchet, A. Spallicci, and B. Whiting, Fundamental Theories of Physics Vol. 162 (Springer, New York, 2011), p. 87.
[54] J. D. Brown and J. W. York, Quasilocal energy and conserved charges derived from the gravitational action, Phys. Rev. D **47**, 1407 (1993).
[55] R. J. Epp, R. B. Mann, and P. L. McGrath, Rigid motion revisited: Rigid quasilocal frames, Classical Quantum Gravity **26**, 035015 (2009).
[56] P. L. McGrath, R. J. Epp, and R. B. Mann, Quasilocal conservation laws: Why we need them, Classical Quantum Gravity **29**, 215012 (2012).
[57] R. J. Epp, R. B. Mann, and P. L. McGrath, On the existence and utility of rigid quasilocal frames, in *Classical and Quantum Gravity: Theory, Analysis and Applications*, edited by V. R. Frignanni, Physics Research and Technology (Nova Science, New York, 2012), p. 559.
[58] P. L. McGrath, M. Chanona, R. J. Epp, M. J. Koop, and R. B. Mann, Post-Newtonian conservation laws in rigid quasilocal frames, Classical Quantum Gravity **31**, 095006 (2014).
[59] P. McGrath, Rigid quasilocal frames, Doctoral thesis, University of Waterloo, 2014.
[60] M. Oltean, R. J. Epp, P. L. McGrath, and R. B. Mann, Geoids in general relativity: Geoid quasilocal frames, Classical Quantum Gravity **33**, 105001 (2016).
[61] M. Oltean, R. J. Epp, P. L. McGrath, and R. B. Mann, Geoids in general relativity: Geoid quasilocal frames, in *The Fourteenth Marcel Grossmann Meeting on General Relativity*, edited by M. Bianchi, R. T. Jantzen, and R. Ruffini (World Scientific, Singapore, 2017), p. 3682.
[62] E. Dewan and M. Beran, Note on stress effects due to relativistic contraction, Am. J. Phys. **27**, 517 (1959).
[63] J. S. Bell, How to teach special relativity, Prog. Sci. Cult. **1**, 1 (1976).
[64] L. Freidel, Gravitational energy, local holography and non-equilibrium thermodynamics, Classical Quantum Gravity **32**, 055005 (2015).
[65] L. Freidel and Y. Yokokura, Non-equilibrium thermodynamics of gravitational screens, Classical Quantum Gravity **32**, 215002 (2015).
[66] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge University Press, New York, 1975).

[67] R. M. Wald, *General Relativity* (University of Chicago, Chicago, 1984).
[68] B. Felsager, *Geometry, Particles, and Fields* (Springer, New York, 2012).
[69] J. M. Lee, *Introduction to Smooth Manifolds* (Springer, New York, 2002).
[70] S. Carroll, *Spacetime and Geometry: An Introduction to General Relativity* (Pearson, San Francisco, 2003).
[71] J. D. Brown, S. R. Lau, and J. W. York, Action and energy of the gravitational field, Ann. Phys. (Amsterdam) **297**, 175 (2002).
[72] J. W. York, Role of Conformal Three-Geometry in the Dynamics of Gravitation, Phys. Rev. Lett. **28**, 1082 (1972).
[73] G. W. Gibbons and S. W. Hawking, Action integrals and partition functions in quantum gravity, Phys. Rev. D **15**, 2752 (1977).
[74] T. Padmanabhan, *Gravitation: Foundations and Frontiers* (Cambridge University Press, Cambridge, England, 2010).
[75] M. Bruni, S. Matarrese, S. Mollerach, and S. Sonego, Perturbations of spacetime: Gauge transformations and gauge invariance at second order and beyond, Classical Quantum Gravity **14**, 2585 (1997).
[76] R. Abraham, J. E. Marsden, and T. Ratiu, *Manifolds, Tensor Analysis, and Applications*, 3rd ed., Applied Mathematical Sciences (Springer-Verlag, New York, 2001).
[77] A. Pound, Self-consistent gravitational self-force, Phys. Rev. D **81**, 024023 (2010).
[78] A. Pound, Gauge and motion in perturbation theory, Phys. Rev. D **92**, 044021 (2015).
[79] D. Puetzfeld and Y. N. Obukhov, Generalized deviation equation and determination of the curvature in general relativity, Phys. Rev. D **93**, 044073 (2016).
[80] A. D. A. M. Spallicci and P. Ritter, A fully relativistic radial fall, Int. J. Geom. Methods Mod. Phys. **11**, 1450090 (2014).
[81] P. Ritter, S. Aoudia, A. D. A. M. Spallicci, and S. Cordier, Indirect (source-free) integration method. II. Self-force consistent radial fall, Int. J. Geom. Methods Mod. Phys. **13**, 1650019 (2016).
[82] S. E. Gralla, A. I. Harte, and R. M. Wald, Rigorous derivation of electromagnetic self-force, Phys. Rev. D **80**, 024031 (2009).